

Rice must be perfectly cooked

The draft genome sequences of *Oryza sativa*, published last week, are impressive achievements, but not the finished article. The public rice genome project more than ever deserves support.

If a base pair of DNA were a grain of rice, then a perfect sequence of the crop's genome would lay out 450 million grains end to end over more than 2,000 kilometres, with each grain correctly aligned and positioned. Assembling this sequence is the goal of the Japanese-led public International Rice Genome Sequencing Project (IRGSP), which is working on the *japonica* subspecies.

But the IRGSP is not the only game in town. The draft genomes of *japonica* by Syngenta, the Swiss-based agribiotech company, and of the *indica* subspecies by the Beijing Genomics Institute, published in *Science* last week, have identified almost all of the rice genes, and will be of immense value to plant breeders. And from the tone of some of last week's media coverage, one could be forgiven for thinking that the rice genome is now 'done' — leaving us standing on the brink of an agricultural revolution, and making the IRGSP redundant.

That would be an unfortunate perception. Funding for the IRGSP runs out in September, and it will need a further US\$10–15 million from its sponsors to finish the job. Policy-makers must understand that the Syngenta and Beijing sequences are rough drafts, with many gaps — more than 130,000, in the case of the Beijing draft — and entire chunks positioned or aligned incorrectly. While better organized than a rice stir-fry, they're still some way from that perfect chain of neatly aligned grains.

The newly proposed collaboration between the private- and public-sector projects (see page 573) means that the money required to finish the rice genome would represent excellent value — rather than completing the project by 2005, it could be done by next year. And

for the same price as is currently budgeted for finishing *japonica*, *indica* could be completed as well. The gains would come not just from sharing sequence data, but also from collaboration on strategies for assembling the sequenced fragments.

Having a sequence with gaps closed and errors corrected is important, for only then may it be possible to fully unravel the complex networks of genes and proteins that govern the crop's biochemistry and anatomy. Once in hand, moreover, this knowledge will transfer to other cereals, which, although they have much larger genomes, have similar genes arranged in much the same order. If the genetics of a particular biochemical pathway is understood in rice, this information can be used to isolate counterparts in other species. A finished rice genome, therefore, will boost efforts to genetically improve a variety of crops — including the attempt to increase yields of rice by giving it the photosynthetic efficiency of maize (see page 576).

But before we get carried away with dreams of feeding the world through transgenic technology, here's a shot of political and economic reality. The availability of the rice genome could make a major contribution to improving food security in the twenty-first century, but ridding the world of hunger and poverty will require more than higher crop yields. Indeed, the impact of the \$10–15 million needed to complete the rice genome needs to be considered in the context of the \$1 billion that rich countries spend daily on subsidies and other measures to protect their farming industries against exports from the developing world. Now there's food for thought. ■

A climate for change

US climate research faces a revamp that may provide political leaders with better scientific information. Will they listen?

The threat posed by global warming was not a top priority for President George W. Bush at the time of his election and appears to be even less of one now. Since his administration announced its intention of walking away from the Kyoto Protocol more than a year ago, there's been little sign of a strategy to take its place. And the failure to re-nominate the respected atmospheric scientist Robert Watson as chair of the Intergovernmental Panel on Climate Change (see *Nature* **416**, 251; 2002) has reinforced the impression of neglect.

Yet there is considerable concern about global warming among the US public. Recent news of the break-up of the Larsen B ice shelf on the Antarctic Peninsula, and the remarkably dry and mild winter that has just ended in North America, have added to this unease. As elsewhere in the world, however, the feeling that global warming is a real problem is held widely, but not particularly deeply. It is not yet an issue on which elections are won or lost.

The one thing that the Bush administration has promised to do is to strengthen the government's research programmes into the scientific understanding of the problem and the technological means available to address it. John Marburger, Bush's science adviser, is

orchestrating a review of the US Global Climate Change Research Program, which was set up by Bush's father to coordinate this research (see page 571).

This attention is overdue. The programme office has never been powerful enough to coordinate the research activities of stridently independent agencies of the US government. As a result, for example, the United States' programme in climate modelling is much weaker than its considerable investment in climate change research should allow. Modellers need supercomputer capacity but lack access to it if they happen to work for the wrong agencies. No wonder that some are looking enviously at Japan's new Earth Simulator facility, and hoping they will be given time on the machine (see page 579).

Care must be taken to ensure that the new mechanisms erected by Marburger result in a balanced programme that is not obsessed with providing short-term answers and dubious technological fixes. But if this mistake is avoided, the revamp could give US climate-change research a new sense of direction, as well as a voice at the most senior levels of the administration. Whether anyone will listen, however, remains to be seen. ■





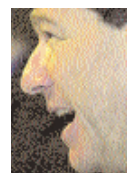
Fish fingered
Fugitives from salmon farms seen as threat to natural stocks
p571



Open access
Restrictions eased over US satellite research
p572



Rice splice
Public and private genome efforts consider a merger
p573



Economic move
Jeffrey Sachs comes down to Earth in New York
p574

Ranches blamed over spread of mad deer

Jonathan Knight, San Francisco

A relative of mad cow disease that afflicts elk and deer appears to be defying efforts to stop its spread across North America.

Some wildlife officials suspect the proliferation of elk and deer ranches is partly to blame for the recent appearance of the disease in new locations in the United States and Canada. Others fear that wild populations in more places than had previously been realized may be harbouring the disease.

Chronic wasting disease (CWD) — a transmissible spongiform encephalopathy, like mad cow disease and the human form Creutzfeldt-Jakob disease — has been endemic among wild populations of deer and elk in areas near where the borders of Colorado, Wyoming and Nebraska meet, at least since the 1960s, when the first cases were discovered.

Although there is no evidence that CWD has been transmitted to either humans or cattle, wildlife agencies are watching the spread of the disease with some trepidation, given the public alarm that any such transmission would generate.

In 1996, CWD was found on an elk ranch in Saskatchewan in Canada, and since then

the disease has turned up on more than a dozen ranches as far south as Oklahoma. All the infected herds have been destroyed, along with herds on farms identified as having received elk from them. Last October, after the disease appeared in herds belonging to one of the largest suppliers of elk in Colorado, the US Department of Agriculture authorized \$2.6 million in emergency spending to keep the disease under surveillance and to compensate affected ranchers.

But in February, routine screening of white-tail deer killed by hunters turned up three infected animals in central Wisconsin, 800 miles east of the closest previously known case. Wildlife officials immediately began culling 500 more deer from the area for testing. By 5 April, nine more cases had been found with about half the testing done.

Lying in the opposite direction from the region where the disease is endemic, the Rocky Mountains had been considered a natural barrier to its westward spread. But in late March and early April, Colorado's Division of Wildlife found two infected mule deer west of the mountains.

Although the deer were wild, they were within the confines of a fence erected in July



The recent discovery of chronic wasting disease in Wisconsin deer has prompted a major cull.

A. MANIS/AP

around a 1,500-acre ranch. Farm-raised elk brought to the ranch in August are a possible source of infection, although that would require a surprisingly rapid transmission, says Todd Malmsbury, chief spokesman for the division. Tests now under way on the elk

Smithsonian rocked by high-level departure

Josette Chen

Unease among researchers at the Smithsonian Institution in Washington DC, the world's largest museum complex, deepened last week with the announcement that Dennis O'Connor, its undersecretary for science, will be departing.

O'Connor, who is also the acting director of the National Museum of Natural History, will become vice-president of research at the University of Maryland at College Park.

The departure adds to the atmosphere of turmoil at the museum complex, which has endured a bitter and public struggle between its cash-starved scientists and its secretary, Larry Small.

A science commission, chaired by Jeremy Sabloff, director of the University of Pennsylvania Museum of Archaeology and Anthropology in Philadelphia, was formed last year to advise on the restructuring of research at the Smithsonian, and will issue recommendations this autumn.

O'Connor says that he does not think his exit will leave the Smithsonian's research in the lurch. He says his decision is not a vote of no confidence in Small, but rather "a statement about College Park and the remarkable opportunity there".

However, Paul Forman, a curator for physics at the National Museum of American History, has branded the latest departure "a clear indication of the

ruination of the institution that the secretary is pursuing".

Small himself was travelling and unavailable for comment. But a member of the Smithsonian's governing board of regents says that Small retains the board's confidence. "He has done the job the regents set out for him," says Manuel Ibanez, a biologist and former president of Texas A&M University-Kingsville.

Ira Rubinoff, director of the Smithsonian Tropical Research Institute in Panama, has been proposed as a possible successor to O'Connor. But Rubinoff has said he would not take the job on an acting basis, and may be considered too close to retirement to be offered it permanently.

and on 300 deer culled from outside the fence should reveal whether CWD has truly breached the Rockies.

Ranching has almost certainly contributed to the spread of CWD. The elk industry exploded in the early 1990s when the animal's velvet antlers, valued in traditional Asian medicine, began to attract prices of up to \$200 a kilogram on the international market. Ranchers in the United States and Canada then began to keep the animals in large numbers. Long-distance rapid transport of commercial elk is speeding the spread of a disease that moved slowly among wild animals, Malmsbury says.

Because the only definitive test for CWD is an examination of the brain stem, surveillance programmes rely on samples from slaughtered or hunted animals, and infected herds are often spotted only after some members have been shipped away. Although it is possible to trace sales of live elk, some states have only recently started requiring ranchers to keep such records. Other states, such as Iowa, operate no surveillance of CWD.

In theory, CWD could be present in the wild at low levels outside the known endemic area without being noticed, says Sam Holland, chief veterinarian of South Dakota. Ranching might bring it to the fore because the disease would spread faster in a concentrated herd. "I don't think anyone knows whether it exists at a low level in populations out there or not," he says. ■

Disbelief greets claim for creation of first human clone

Alison Abbott

European scientists have voiced scepticism about claims by the Italian gynaecologist Severino Antinori that one of his patients is two-months pregnant with a cloned human embryo.

Antinori, who is director of the infertility unit at the International Associated Research Institute for Human Reproduction in Rome, made the announcement to the newspaper *Gulf News* last week, when he attended a meeting on human genetics in the United Arab Emirates.

One Italian newspaper, *Il Tempo*, said that Antinori had confirmed the claim, but he has refused to comment to other journalists, saying only that "research needs silence". Glasgow's *Sunday Herald*, meanwhile, reported that Antinori had split with his long-time collaborator, Panos Zavos of the Andrology Institute of America in Lexington, Kentucky, and that Zavos was pursuing his own cloning programme independently in Cyprus.

"There is, of course, some scepticism that the cloning has really been done," says Carlo Redi, an animal-cloning expert at the University of Pavia. Although the work is alleged to have been done outside Europe, if it has taken place, he says, it contravenes both the

law in Italy and normal codes of medical practice. Studies in animals have shown that cloning yields few successful births, and that a high proportion of the offspring produced are born with defects. The Italian Medical Association has issued several threats to expel Antinori, but has not yet done so.

Caution is also apparent outside Italy. "It is theoretically possible," says Eckhard Wolf, a cloning expert from the University of Munich, but he adds that "even if Antinori has achieved a pregnancy, the rate of spontaneous abortion in the second trimester of pregnancy is very high, so comments are premature."

More bluntly, Ian Wilmut of Edinburgh's Roslin Institute, a member of the team that cloned Dolly the sheep, says that Antinori's claim is "either a misunderstanding or deliberately misleading". The biggest danger, he says, is that Antinori's activities will generate public anxiety about cloning, turning people against the idea of using 'therapeutic cloning' for regenerative medicine.

In Italy, scientists fear that the news could adversely affect the public mood just as a bill to regulate the use of embryonic stem cells and of therapeutic medicine is to be debated in parliament. "We worry that it may increase pressure for a very restrictive law," says Redi. ■

Court rules out Italian research appointment

Sally Goodman

Italy's national research council, the CNR, has been found guilty of unfair treatment towards an Italian astrophysicist during the selection process for a directorship at one of its institutes.

The court ruling could have wide implications for the other appointments so far made in the CNR's international recruitment drive to find leaders for 100 of its institutes (see *Nature* 414, 133; 2001).

Giovanni Bignami, currently research director of the Italian space agency, ASI, was interviewed for the directorship of the newly created Institute of Astrophysics and Cosmic Physics last October. But the post was awarded to the director of the Institute for Extraterrestrial Radiation in Bologna, one of four centres making up the new astrophysics institute.

Bignami challenged the appointment in a Rome court on the grounds that the CNR had not followed the correct procedure during the selection process. He argued that the research agency had failed to make sure that the nine-member executive council of

the CNR, which makes such appointments, received proper advice from specialists on the candidates' qualifications.

"It was a blatant case of injustice," says Bignami. The court decision, issued on 4 April, overturns last autumn's appointment and will probably force the CNR to rerun the recruitment process.

Reactions to the judgement were mixed. Arturo Falaschi director of the International Centre for Genetic Engineering and Biotechnology in Trieste, a member of the council who voted for Bignami, says that the court's judgement sets a "terrible precedent". He insists that the CNR went out of its way to make the recruitment process public and transparent. "I believe Bignami was the best person for the job," he says, "but due process was followed and the final decision was legitimate."

But an adviser to the research and education ministry says that he was unhappy about the way the CNR had handled the appointment and is pleased with the court's judgement. "The CNR and



Guilty: Italy's national research council was unfair when it selected one of its directors.

the ministry should now be questioning why this has happened. Other nominations may now have to be cancelled," the adviser says.

A spokesperson for Lucio Bianco, the CNR's president, said that Bianco would not comment until he had read the full text of the judgement which, as *Nature* went to press, had yet to be released by the court. ■



Food for thought: publicity over the Orkney fish breakout has raised the profile of ecological concerns.

Stream of escaped farm fish raises fears for wild salmon

Natasha McDowell, London

The escape of an estimated 100,000 farmed salmon in the Orkney Islands, off the north coast of Scotland, has highlighted mounting concerns about the ecological impact of such incidents on natural salmon stocks.

A Scottish parliamentary committee is currently investigating the impact of fish farming. But the Scottish Executive, which has been responsible for fishing and the environment since its establishment in 1999, has been accused of ignoring the issue.

To judge from scenes in Orkney last week, when hundreds of locals appeared to relive the movie *Whisky Galore* as they scrambled to harvest the salmon, many of the escaped fish will end up safely ensconced in home freezers.

But according to data collected in Norway and elsewhere, farmed fish from less celebrated escapes are contaminating and sometimes overwhelming the gene pools of natural fish. The offspring of farmed fish, some data suggest, are often unable to complete the heroic salmon runs by which the natural species navigate between spawning grounds inland and breeding grounds in the ocean. Critics say that, together with the rampant transmission of lice and disease from fish farms to natural stocks, the result is threatening the very survival of natural salmon runs in countries such as Scotland, Canada and Norway.

Kenny Black, a marine biologist at the Dunstaffnage Marine Laboratory in Oban, has been asked by the Scottish parliament to prepare a technical report for its inquiry. "The increasing level of escapes means we risk the extinction of the genetic integrity of native salmon populations," he says. "Farmed fish are fairly homogenous and so there will be an irreplaceable loss of genetic diversity." However, as Black notes, "fish farming is here to stay. What we need to do is identify the areas of most concern and do something about them."

Black's report, when it is delivered to the

parliament's committee on transport and the environment next month, will indicate where more research is needed to understand better the impact of escaped farm fish.

Although containment techniques for farm fish have improved, the number of escapees has not fallen because the industry is expanding, says Black. According to environmental groups, about a million salmon have escaped from farms in Scotland since 1998.

"So far, the government has done nothing," says Robin Harper, a Green member of the parliament.

An initial report from the committee, released last month, called for tagging to allow escaped fish to be traced, and the introduction of associated penalties for fish farmers. It also recommends that farm fish be sterilized.

Ian Fleming, a marine biologist at Oregon State University, who has studied the impact of salmon farming in Norway, says that official counts of escaped farm fish there are "conservative estimates". As a result of the escapes, he says, differences between wild and farmed populations are halving every 10 generations. In some rivers 80% of salmon are of farmed origin.

Fleming has found that, where farmed fish have mixed with natural ones, 30% less offspring make it from breeding grounds to the ocean. He says that the negative effects of hybridization do not usually become apparent until the second generation, when the co-adapted genes that tell the fish when to breed and how to find food become separated.

Research on the mortality rates of second-generation fish is being undertaken by Andrew Ferguson of Queen's University Belfast, and will be published shortly. The results could lead to new rules for fish farming. "We need a balanced approach soundly based on scientific evidence," says Bristow Muldoon, a Labour member of parliament and chair of the Transport and the Environment Committee. ■

White House plans to increase 'relevance' of climate research

Mark Schroppe

The United States is planning to restructure its climate-change research programmes in an effort to make them more relevant to government policy.

White House officials say that the reorganization will help the government to answer scientific questions that influence policy, as well as improving coordination of the entire research effort.

At the moment, climate-change research is coordinated by the US Global Change Research Program (USGCRP), which was set up in 1989 by former President George Bush.

But officials in the current administration say that the USGCRP — which is overseen by one of seven government committees that deal with technical issues — is too far removed from cabinet-level decisions. They also say that its scope excludes large chunks of climate-change research, including, for example, about half of the related activities at the National Oceanic and Atmospheric Administration.

Under the proposals, a new Committee on Climate Change Science and Technology Integration, whose members will include most members of the cabinet, will be responsible for making recommendations to the president on climate-change issues. In cooperation with a working group on climate-change science and technology, the new committee will set research priorities for climate change, White House documents say.

Two new offices, a Climate Change Science Program Office and a Climate Change Technology Program Office, will coordinate the government's climate-change research.

John Marburger, director of the White House Office of Science and Technology Policy, will help to run the integration committee. He says that the new structure will increase the impact of climate-change research on policy by focusing research on the government's priorities.

But Richard Moss, director of the USGCRP coordination office, says that it will be a challenge for people carrying out long-term basic research to compete with those in areas that have potential for short-term policy application.

Marburger aims to have the system running by the end of the year, and does not envisage problems along the way. "I don't think there are too many different ways of accomplishing what's needed," he says. ■

Chinese institute is an innovation in theory



Mould-breaking: the Chinese Academy of Sciences is going multidisciplinary at its new institute.

David Cyranoski, Shanghai

Interdisciplinary boundaries are rarely stricter than those that separate the 100-plus research institutes of the Chinese Academy of Sciences, home to some 40,000 of the country's researchers. But the academy is moving tentatively towards breaking that mould by setting up its first multidisciplinary institute.

The Shanghai-based Institute for Advanced Studies will eventually employ 20 research fellows from various disciplines and will tackle complex problems in biology, concentrating on their theoretical aspects.

"This is essential for China," says Uli Schwarz, a biologist at Germany's Max Planck Institute for Developmental Biology in Tübingen and a co-director of the new institute. "We need to create a new atmos-

phere here," adds Schwarz, who will spend around a third of his time at the institute.

In an effort to bolster its scientific ties with China, Germany's Max Planck Society and other German sources will be providing part of the institute's US\$800,000 annual running costs.

Researchers are expected to spend between one month and a year at the institute. Officials promise that they will be free to follow their research in any direction. Those selected will specialize in fields of biology with a strong theoretical component, such as evolutionary biology, and will also investigate the interactions between science and society.

Organizers hope that research into science-based regulations, such as controls on pollution, and the history of science, will enable the institute to provide a broad-based view of the challenges faced by the Chinese Academy. "As scientists, we have to take note of the effect we are having on society," says Schwarz.

Fellows of the institute will be expected to publish their ideas on these subjects. If they need to do bench work, arrangements will be made for this to be carried out at the Shanghai Institute of Biological Sciences, says Yi Rao, a molecular neurobiologist at Washington University in St Louis and the institute's other co-director.

The institute held its first conference at the end of last month, on the theory of consciousness. International participants included Patricia Goldman-Rakic, a neurobiologist at Yale University School of Medicine; psychologist Willem Levelt from the Max Planck Institute for Psycholinguistics at Nijmegen in the Netherlands; Christof Koch, a neuroscientist at the California Institute of Technology; and philosopher Thomas Metzinger of the Johannes Gutenberg University in Mainz, Germany.

"The conference was a great success in linking psychological and philosophical perspectives with well-grounded neurological science," says David Van Essen of Washington University in St Louis.

A second conference, on social behaviour, is to be held in September, when the institute plans to move to its permanent home on the campus of the Shanghai Institute of Biological Sciences, which is also part of the Chinese Academy of Sciences. The institute is already searching for further sources of funding, both inside and outside China.

Speaking privately, some participants at last month's meeting questioned whether an institute that will concentrate on theoretical questions will be able to maintain support from a government that is increasingly giving priority to applied science.

US strikes a deal on export rules

Tony Reichhardt, Washington

The US government has compromised on export restrictions applying to researchers building satellites and other space hardware. But space scientists will still face obstacles in working with foreign colleagues, experts say.

The State Department issued an interim rule on its International Traffic in Arms Regulations (ITAR) on 29 March. Since 1999, satellites have been under the jurisdiction of ITAR. Scientists have grown concerned about their possible liability under the rules, which can require licensing not only for the export of satellites and spacecraft instruments, but also for technical discussions about them. As a result, some foreign-born researchers at US universities have been barred from fully participating in international space missions (see *Nature* 404, 321; 2000).

University lobby groups had pushed for a blanket ITAR exemption for scientists engaged in "fundamental research" where information is in the public domain. The new rule is "certainly not everything the universities wanted", according to Tim Brightbill, an attorney with law firm Wiley Rein & Fielding, who advises academic and commercial clients on ITAR compliance.

Under the new rule, US universities will not need ITAR licences if they export satellite hardware or information to members of NATO, the European Union or the European Space Agency, and countries listed as "major non-NATO allies", such as Japan and Israel. But they still need a licence for exports and exchange of technical information with other foreign nationals, including researchers from India or China. And collaborators in approved countries

would have to guarantee that no researchers from unapproved countries were receiving ITAR-restricted information.

Supporters of stiff export controls, who opposed NASA in negotiations over how much to ease the ITAR restrictions, are unlikely to concede any more ground, says Jack Shane, also of Wiley Rein & Fielding. Sensitivity about terrorism will weaken the case of those who favour open access.

The State Department issued the regulation as an 'interim final rule', however, leaving the door open for additional public comment on it.



Feeling blue: projects such as Gravity Probe-B have been hurt by export-control rules.

Rice genome sequencers cook up merger

Declan Butler

The two public groups and two private companies sequencing the rice genome are discussing ways to merge their efforts into a single joint project.

Researchers believe that such cooperation would greatly accelerate the completion of the genome. Draft maps for two subspecies of rice (*Oryza sativa*) were published on 5 April — Swiss-based agricultural biotechnology company Syngenta covered *japonica* (S. A. Goff *et al. Science* **296**, 92–100; 2002), and the publicly funded Beijing Genomics Institute sequenced *indica* (J. Yu *et al. Science* **296**, 79–92; 2002).

Representatives of these groups, together with those from biotechnology firm Monsanto of St Louis, Missouri, and from the other public project, the Japanese-led International Rice Genome Sequencing Project (IRGSP), could meet within weeks to organize the merger. If they reach agreement, an announcement of the deal could take place at next month's IRGSP workshop in France.

The move follows complaints from some researchers about access to sequence data for Syngenta's draft map of *japonica* (see *Nature* **416**, 111–112; 2002). Syngenta makes the data available to academic researchers for non-commercial purposes at its own website, rather than depositing the information in GenBank.

But access is governed by a number of conditions. Researchers can only search up



Steaming ahead: cooperation would speed up the completion of the rice genome sequence.

to 15,000 base pairs of the sequence at a time, and download only 100,000 base pairs of sequence each week. More can be freely downloaded if the researchers' institutions sign an agreement with Syngenta stating that the data will not be used for commercial ends.

But Syngenta last week announced that it will make its data available to the IRGSP, effectively ensuring that they end up in GenBank, mingled with the IRGSP's results. Syngenta completed its draft sequence last year but, despite providing the IRGSP with mapping data, it had until now refused to share raw sequence data with all members of the public project — unlike its rival Monsanto, whose contributions of sequence data are credited with accelerating the IRGSP's effort.

The merger idea came in a written proposal to the IRGSP from Stephen Goff, head of Syngenta's Torrey Mesa Research Institute in San Diego, California, and Jun Yu, associate director of the Beijing Genomics Institute.

The proposal would mirror the cooperation between Celera, the Maryland-based genome-sequencing company, and the public project headed by Gerry Rubin at the University of California, Berkeley, on the fruitfly *Drosophila*. That deal sped up assembly of the genome and improved data quality (see *Nature* **401**, 729–730; 1999).

Cooperation on rice could result in the genome being completed next year — two years ahead of the current schedule, participants in the discussions say. The collaboration might also extend to annotating the raw sequence with information such as the identity of genes and their function in rice. ■

portal.tmri.org/rice

MONSANTO

Government warned over future of cash-starved labs

David Adam, London

Out of date, out of shape and in need of heavy investment, the laboratories of Britain's universities have joined its railways and health service as a cause of national lament after a damning report found they urgently need a £3-billion (US\$4.3-billion) fix.

The report, issued on 25 March, was commissioned by the Office of Science and Technology (OST) from consultancy firm JM Consulting. It says the lack of investment in buildings, support staff and equipment could threaten Britain's scientific future.

"Without this investment there will remain a serious problem, which is causing significant stress to institutions, and actual damage to the present and future research output," the report warns. "There is the expectation of more serious and destabilising effects if current trends continue." The OST will use the report, *Study of Science Research Infrastructure*, to try to squeeze more money out of the forthcoming governmental spending review.

The study says that recent government initiatives such as the Joint Infrastructure Fund (JIF) and the Science Research Infrastructure Fund (SRIF) — which offered universities £2 billion to improve facilities — have had little impact. Only 15% of labs got any money, as most of it was went into construction projects with tangible research benefits, rather than general maintenance.

"We now have a new building with capacity for 450 scientists but no money for a porter or a receptionist," said an official at one of the 20 unnamed institutions contacted during the survey.

JIF required universities to match any funds provided, and many institutions did this by draining the cash from other projects.

"The ability to 'make do' has been a powerful factor to the benefit of UK science," the report says. However, a new generation of researchers is less tolerant of "working long hours in poor environments with inefficient equipment", it says.

But lobbyists are reluctant to attack a

government that has improved science funding in recent years. "It is not realistic to expect years of underinvestment to be reversed overnight," says Richard Joyner, chairman of pressure group Save British Science. "JIF and SRIF were just the first attempt to sort out a very difficult problem." ■



Australia gives a cautious approval to stem-cell research

Sydney Australia's political leaders have given a qualified go-ahead to stem-cell research. Scientists will be allowed to extract stem cells from surplus embryos that were created by *in vitro* fertilization before the decision was made on 5 April, but will not be allowed to create cloned embryos for reproductive or therapeutic purposes.

The decision, which follows months of public debate and intense lobbying by scientists (see *Nature* **416**, 4; 2002), was made at a meeting of the Council of Australian Governments, a body composed of federal, state and territorial leaders. The move means that Australia's researchers will have access to the estimated 70,000 embryos that are held in frozen storage in the country's fertility clinics, subject to approval from regulatory bodies and the consent of the donors.

The decision not to use embryos created after 5 April was taken after some politicians warned that demand for embryos could lead to the development of an 'embryo farming' industry. The ban will be reviewed by an ethics committee after 12 months, and the ban on therapeutic cloning will be assessed after three years.

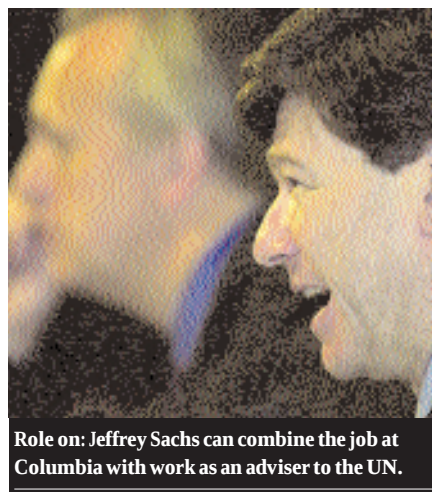
Academics urge Europe to suspend funds for Israel

London Israel should be prevented from receiving funds from European research organizations until it opens peace negotiations with the Palestinians, a group of European scientists has said.

A total of 125 academics — including Patrick Bateson, provost of King's College at the University of Cambridge, and Giorgio Parisi, a physicist at the University of Rome La Sapienza — made the call in a letter published in *The Guardian* on 6 April.

The scientists say that many European funding institutions, such as the European Union and the European Science Foundation, regard Israel as a European state when it comes to awarding grants and contracts. They propose a moratorium on research funding for Israel until it abides by UN resolutions on Palestine and begins peace negotiations.

But a moratorium could damage research projects that were designed to promote collaboration between Israeli and Palestinian scientists. Economic sanctions against Israel were proposed at a meeting of European foreign ministers in Luxembourg last week, but European Union officials say that research funding was not discussed.



Role on: Jeffrey Sachs can combine the job at Columbia with work as an adviser to the UN.

Top economist quits Harvard for New York

New York Economist Jeffrey Sachs is leaving Harvard University after 22 years to become director of the Earth Institute at Columbia University in New York.

The institute, which employs 1,000 scientists and has a budget of around US\$100 million, currently focuses on natural sciences. Columbia officials say that Sachs will help to raise the profile of the institute's research in social science and economic development.

Sachs says that he was persuaded to move by the chance to combine his academic work with his other new position, that of special adviser to United Nations secretary general Kofi Annan. "There's only one place in the world where one could do that — at Columbia with the UN down the block," he says.

Institutes seek common knowledge on rare diseases

Paris Europe is turning its attention to rare diseases with the instigation of three national research projects.

France is establishing a 'virtual institute' for rare diseases using 1.5 million euros (US\$1.3 million) from the government and the country's Association for Neuromuscular Disease. The institute will be headed by Alain Fischer of the Necker Hospital for Sick Children in Paris, who in 2000 successfully used gene therapy to treat children suffering from a rare defect of the immune system.

In Germany, 10 research projects into rare diseases are scheduled to get under way this December, and the Spanish government is also planning a rare-disease institute.

The European Parliament defines rare diseases as those that will affect less than 5 in 10,000 people during their lifetimes. Between 5,000 and 8,000 such conditions are thought to exist, of which about 80% are genetic in origin. They are often overlooked by drug companies, although both the

European Union and the United States encourage firms to invest in such treatments, for example by allowing the companies to retain exclusive rights to the drugs for longer than normal.

Grave error preserves Stone Age skeletons

Munich An untouched Stone Age communal grave containing at least 40 human skeletons has been unearthed in eastern Germany. Graves of this type are common in the area, but farming practices destroy most of them before archaeologists can investigate them.

The grave was found by researchers from the State Department of Archaeology in Sachsen-Anhalt, just centimetres below the surface of a field 100 kilometres west of the town of Magdeburg. It is around 30 square metres in size. The researchers now plan to use DNA analysis to determine the family relationships between the bodies, which are believed to date from between 3400 and 2800 BC. They will also study the bones to determine the causes of death and the professions of the deceased.

Excavations at the site began a year ago, but the grave was not found until last week because the researchers had built a temporary hut directly on top of it.

France heads for harmony in education by degrees

Paris French students should find it easier to move between European institutes in the future, thanks to a reform of the country's system for granting higher degrees.

The current range of postgraduate degrees awarded by the different components of France's higher education system confuses foreign institutes, officials say, and hampers the mobility of the country's researchers. Education minister Jack Lang announced last week that two of these diplomas — the DEA and DESS degrees — will, after September, start to be replaced with standard two-year masters qualifications.

The move brings France closer to the system of three-year bachelors, two-year masters and three-year doctoral degrees that European countries have agreed to move towards. Italy reformed its higher education system last year, introducing the new system to replace its traditional Laurea degree, and Germany and the Netherlands have also begun to bring their systems into line.

Press centre speaks up for the media shy

London Help is at hand for British researchers who find it difficult to deal with the press.

The Science Media Centre, which aims to help working scientists to communicate their work to the public, opened its doors

in London last week. The centre is based at the Royal Institution, a body primarily devoted to science communication. It plans to make available a 'science spokesperson' to provide comment on controversial areas, such as the farming of genetically modified crops.

"On the one hand we have a public with an apparently poor grasp of the way science works, and on the other we have many scientists who are equally poor at engaging with the media," says Fiona Fox, head of the centre. Fox also hopes that the centre's location in central London will help it to become an established venue for press conferences, briefings and radio and television interviews.

► www.sciencemediacentre.org

Blood-clot study takes off with offer of free air miles

London British and South African scientists are hoping that an offer of 5,000 free air miles will entice volunteers for what they say is the largest study to date on whether long-haul flyers risk developing blood clots in their legs.

The study has been christened BEST — Business class versus Economy Syndrome as a cause of Thrombosis — and will address three questions: whether long-haul air travel affects clotting, the role of existing blood disorders in the process, and whether passengers in cramped economy-class seats are at more risk than those in business class. The project is being undertaken by teams from King's College London and the University of the Witwatersrand in South Africa, and is backed by South African Airways, which has agreed to provide the air miles.

The scientists hope to recruit 1,000 volunteers travelling from Heathrow to Johannesburg between 9 April and 9 May. Blood samples will be taken before and after the flight, and subjects will receive an ultrasound scan of their legs on arrival at Johannesburg airport.



Plane pain: cramped conditions on long-haul flights have been linked with thrombosis.

The rice squad

Feeding the world in the twenty-first century could require a second green revolution. But that may involve the most audacious feat of genetic engineering yet attempted, says Christopher Surridge.

First, some sobering facts and figures: over the next 50 years, the world's population is expected to increase by a third, to some 8 billion. Much of this rise will be in southeast Asia, where half the population is already poorly nourished.

Global production of rice (*Oryza sativa*), the world's most important staple crop, has risen threefold over the past three decades — and if similar progress could be maintained, there should be enough food to meet the rising demand. But here's the rub: yields are fast approaching a theoretical limit set by the crop's efficiency in harvesting sunlight and using its energy to make carbohydrates.

So to feed the world, rice may have to be re-engineered at the biochemical level. "The only way to increase yields and reduce the use of nitrogen fertilizers is to increase photosynthetic efficiency," argues John Sheehy, an ecologist at the International Rice Research Institute near Manila, the Philippines, who specializes in modelling crop yields. He calculates that a boost of 20% should do the trick.



Bowled over: John Sheehy believes improving rice photosynthesis is vital for food production.



Reaching a plateau: millions of people rely on rice, but the crop is nearing its limit in efficiency.

Easily said, rather more difficult to do.

The first green revolution, which began in the late 1960s, depended on dwarf varieties of rice, wheat and maize that, put simply, made more grain for less stem. It has since been shown that this relied on breeding mutants that do not produce, or fail to respond to, the plant hormone gibberellic acid.

Grain power

Now we can add the techniques of genetic engineering to the old methods of mutation and traditional plant breeding. The modified crops created so far generally involve the addition of a single gene — for a natural insecticide, say, or resistance to a herbicide. For geneticists, this is like improving a car's performance by adding an aerodynamic spoiler or changing the tyres. But boosting a plant's photosynthetic efficiency will involve manipulating a whole suite of genes. By the same analogy, it's like supercharging a car's engine by fitting a new fuel injection system.

Thankfully, evolution provides plant scientists with precedents. On at least 30 separate occasions¹, different plant lineages have evolved to use the Sun's energy more efficiently, making sugars in a two-stage process known as C_4 photosynthesis. Exactly when this trick first evolved is unknown, but

from about 10 million years ago, falling concentrations of carbon dioxide in the atmosphere gave plants using C_4 photosynthesis an important selective advantage. The ancestors of maize were among these plants. But rice, wheat and most other cereals all use conventional C_3 photosynthesis.

The terminology refers to the number of carbon atoms in the first molecule created when atmospheric CO_2 is assimilated by the plant. In C_3 plants, CO_2 reacts with ribulose 1,5-bisphosphate (RuBP), catalysed by an enzyme called ribulose 1,5-bisphosphate carboxylase-oxygenase (Rubisco). The product is phosphoglycerate, an organic molecule with a backbone of three carbon atoms, which is then used to build sugars and more complex carbohydrates (see diagram, opposite).

In C_4 plants, CO_2 is assimilated using a different enzyme — phosphoenolpyruvate carboxylase (PEPC). This attaches CO_2 to the three-carbon compound phosphoenolpyruvate (PEP) to produce oxaloacetate, a molecule that contains four carbon atoms. Oxaloacetate can be converted into a series of other four-carbon products, including malate, citrate and aspartate. These four-carbon molecules are broken back down into three-carbon compounds and CO_2 , which is then fed to Rubisco.

This two-stage process is more efficient because it allows C_4 plants to boost the levels of CO_2 reaching Rubisco. PEPC has a stronger affinity for CO_2 than Rubisco, and so can work better at low concentrations of the gas. In addition, Rubisco is inhibited by oxygen — an unavoidable product of the light-harvesting stage of photosynthesis — whereas PEPC is unaffected.

C_4 plants shield Rubisco from the debilitating effects of O_2 by packaging the enzyme's activity into specialized cells, isolated from the O_2 -producing reactions of photosynthesis. In maize these cells surround a leaf's veins, and are called bundle sheath cells. PEPC operates happily in 'mesophyll' cells, found in the rest of the leaf, immune to the build up of O_2 . Four-carbon molecules are then shunted across to the bundle sheath cells, where they are broken down to create locally high concentrations of CO_2 . This allows Rubisco to work with maximum efficiency — and has knock-on effects that lessen the waste of nitrogen, used to make amino acids, and of water.

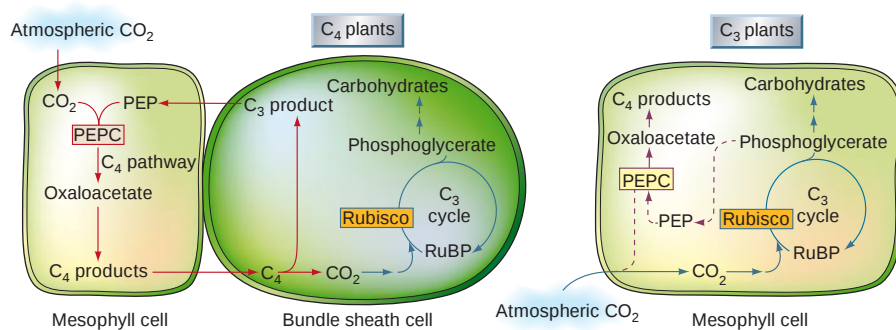
Separate lives

In C_4 plants, the separation of photosynthesis into two pathways operating in different types of cell involves what is dubbed a 'Kranz' anatomy. This is strikingly evident in maize, where alternating bands of mesophyll and bundle sheath cells form stripes running the length of the leaf.

Given that it would involve major changes in both biochemistry and anatomy, turning rice into a typical C_4 plant is no trivial matter. But that hasn't stopped some plant scientists from tinkering with the genes that control photosynthesis reactions in an effort to imbue rice with some C_4 characteristics.

Photosynthesis in maize differs from that in rice in the activity of seven enzymes. Genes for some of these enzymes are present in the *O. sativa* genome — drafts of which were published last week^{2,3}. But some of the rice versions — including that for PEPC, which in rice can make oxaloacetate from the products of the C_3 pathway — may not be suitable for bona fide C_4 photosynthesis. So researchers led by Maurice Ku of Washington State University in Pullman are attempting to introduce key genes from maize into rice.

In 1999, Ku collaborated with two Japanese groups, led by Makoto Matsuoka of Nagoya University and Mitsue Miyao of the National Institute of Agrobiological Sciences in Tsukuba. Together, they used *Agrobacterium tumefaciens* — a bacterium that can insert foreign sequences into the genomes of its hosts — to introduce the maize gene for PEPC, including its regulatory sequences, into two cultivars of rice. They found that transformed plants containing multiple copies of the gene produced two or three times as much PEPC as ordinary maize. Encouragingly, photosynthesis in the transformed plants



Supercharged: two-stage C_4 photosynthesis is more efficient than the C_3 version found in most plants.

was inhibited less by oxygen than in control rice plants that had not taken up the gene⁴.

This was merely the first step. Ku and his colleagues are also working with the maize genes for two other enzymes, pyruvate orthophosphate dikinase (PPDK) and NAD malate decarboxylase. The former is involved in the production of PEP; the latter breaks down four-carbon molecules to feed CO_2 to Rubisco. Ku's team plans to make transgenic rice plants that express each of the three genes, and then interbreed them to incorporate all of the genes into the same plants.

Rice plants expressing either the maize gene for PEPC or PPDK seem to have a photosynthetic capacity about 30% greater than normal plants⁵, and in preliminary field trials have produced 12% and 35% more grains, respectively. Ku's team has also made plants expressing both genes. These have double the photosynthetic capacity of normal rice⁵ — and, in small-scale field trials in China and South America, they have produced up to 90% more grains. "These



Maurice Ku is trying to add maize genes to rice.

plants are also more vigorous and uniform, more tolerant to stress conditions, and flower four to six days earlier than non-transformed controls," claims Ku.

Such results make it sound as if the second green revolution is already in full swing. But most experts urge caution, because the gains seen by Ku and his colleagues may actually have little to do with the acquisition of C_4 photosynthesis — instead, they might be a result of improved stress tolerance. The production of PEPC can increase dramatically in normal rice plants grown under stressful conditions. And the most spectacular gains in yield seen in the field trials came in plants grown in the stressful environment of South American soils, which have a high salt concentration and where ultraviolet wavelengths make up a significant proportion of the sunlight.

Ku and his colleagues also note that the introduced genes increase the opening of pores called stomata on the surface of the

plants' leaves^{4,5} — which would increase the amount of CO_2 available for photosynthesis even in the absence of a functional C_4 pathway. Still missing, say plant scientists, is any biochemical proof that the transgenic plants are concentrating CO_2 at Rubisco using four-carbon molecules. "The evidence is not there yet," says plant physiologist and ecologist Barry Osmond, president of Columbia University's Biosphere 2 Center in Oracle, Arizona. "It comes down to keeping an open mind but not overstating things."

Division of labour

Even if the full complement of C_4 enzymes can be transferred to rice, reaping all the benefits requires PEPC to be kept physically separate from Rubisco to prevent the latter from being inhibited by oxygen — which brings us back to Kranz anatomy. Unfortunately, plant scientists do not yet know how to conjure up this anatomical arrangement. Although the genes for the C_4 photosynthetic enzymes are well known, those responsible for directing the development of Kranz anatomy remain mysterious.

But the latest research suggests that C_4 and C_3 plants are more similar than was once thought — which has boosted hopes of finding a simple genetic switch to convert rice to a Kranz anatomy. In January this year, British researchers reported the discovery of C_4 characteristics hidden in tobacco (*Nicotiana tabacum*), a C_3 plant⁶. Julian Hibberd of the



Could a simple genetic switch make rice capable of meeting the world's food needs?



Sowing the seeds: the International Rice Research Institute is backing the effort to boost rice yields.

▶ University of Cambridge and Paul Quick of the University of Sheffield studied photosynthetic cells surrounding vessels in tobacco stems — the same type of cells that give the veins of celery stalks their prominent green colour. The researchers found active C_4 enzymes, and showed that the cells broke down malate to release CO_2 , just like the bundle sheath cells in maize leaves.

Trigger happy

That even archetypal C_3 plants contain cells functionally equivalent to the bundle sheath cells of maize helps to explain why C_4 photosynthesis has been able to evolve repeatedly. Intriguingly, it also suggests that if plant scientists can find the genetic triggers to induce the cells studied by Hibberd and Quick to develop in rice leaves, it may not even be necessary to engineer in the genes for C_4 enzymes.

There are precedents for finding simple genetic changes that underlie a major anatomical shift. For example, domesticated maize (*Zea mays* ssp. *mays*) arose from a wild relative called teosinte (*Z. mays* ssp. *parviglumis*) in Mexico between 5,000 and 10,000

years ago. Teosinte is a bushy plant with a lot of branching stems, each carrying many male flowers and small female ears. In maize, these side branches have all but disappeared, leaving a single stem with one male flower and several large, grain-bearing cobs. This transformation was achieved largely through changes in a single gene, *teosinte branched1* — mutations in which can restore maize to an ancestral, teosinte-like, appearance⁷.

But what are the genetic switches that have given rise to Kranz anatomy? Studies of maize may again prove useful. Although maize leaves have a well-defined Kranz anatomy and C_4 photosynthesis, other leaf-like organs, such as the husks that wrap around developing cobs, do not. Instead their cells are more homogeneous and use C_3 photosynthesis⁸. Jane Langdale of the University of Oxford suspects that this pattern is the 'ground state', with the C_4 mesophyll and bundle sheath cells being more differentiated variations.

Golden nuggets

Langdale has identified a family of genes, dubbed *G2-like* (*Glk*), that is involved in the development of chloroplasts — the sub-cellular organelles in which photosynthesis occurs. The genes are named after the first member of the family, *Golden2* (*G2*), which was discovered in maize. Their sequences suggest that they encode proteins that act as transcription factors, regulating the expression of other genes⁹.

In maize leaves, *G2* is expressed primarily in bundle sheath cells whereas another member of the family, *ZmGlk1*, is active in mesophyll cells¹⁰. But in maize husks — and in rice leaves — the two main *Glk* genes are expressed at similar levels in all cells. If the differential expression of *Glk* genes is crucial to establishing Kranz anatomy, it should be possible to identify the genes responsible for controlling the *Glk* genes' activity, and to engineer them into rice.

More exotic plants might provide further clues. The viviparous spikerush (*Eleocharis vivipara*), a leafless sedge that lives in the

swamps of Florida, is a C_3 plant when submerged. But when it grows above the water, its stems develop Kranz anatomy and full-scale C_4 photosynthesis^{11,12}. Buried in such amphibious plants might be a master switch to turn on C_4 characteristics.

It may not even be necessary to engineer Kranz anatomy into rice. *Hydrilla verticillata* is an aquatic C_4 plant that keeps PEPC and Rubisco separate, not in different cells, but by operating the former in the cytoplasm of mesophyll cells, and the latter in its chloroplasts¹³. *Borszczowia aralocaspica*, which lives in salty depressions in the semi-deserts of central Asia, performs a similar trick by packaging the two enzymes into different chloroplasts¹⁴.

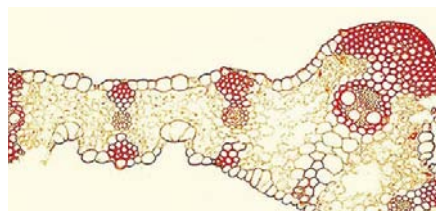
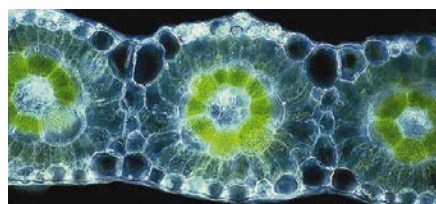
Hydrilla and *Borszczowia* are probably evolutionary halfway houses between typical C_3 and C_4 forms. "This may offer an alternative to engineering C_4 using Kranz anatomy," says Gerald Edwards of Washington State University, who led the team that discovered the compartmentalization in *Borszczowia*. But that still means identifying the genes responsible and engineering rice to develop in a similar fashion.

Given such obstacles, turning rice into a C_4 plant is an endeavour that could take a decade or two. Aside from the scientific challenges, funding could become an obstacle. Japan is the most significant benefactor for rice research, and its largesse will be needed to support efforts to re-engineer the crop's photosynthetic efficiency. But given the present state of the country's economy, there is currently talk of Japan slashing its support for the International Rice Research Institute, which is the natural home for the project.

Many of the non-governmental bodies committed to fighting world hunger, meanwhile, remain vehemently opposed to plant genetic engineering — which they see as a method of enriching multinational companies, rather than a means of achieving food security for developing countries.

But enthusiasts argue that supercharging photosynthesis in rice to achieve a second green revolution might be just the project to change hearts and minds. "This is the next frontier," Sheehy asserts. ■

Christopher Surridge is a senior biology editor at Nature.



C_4 plants have Kranz anatomy (top) in which different photosynthetic pathways take place in bundle sheath cells (bright green) and mesophyll cells, a distinction not seen in C_3 plants (below).

1. Sage, R. F. *Plant Biol.* **3**, 202–213 (2001).
2. Yu, J. et al. *Science* **296**, 79–92 (2002).
3. Goff, S. A. et al. *Science* **296**, 92–100 (2002).
4. Ku, M. S. B. et al. *Nature Biotechnol.* **17**, 76–80 (1999).
5. Ku, M. S. B. et al. *Novartis Found. Symp.* **236**, 100–111 (2001).
6. Hibberd, J. M. & Quick, W. P. *Nature* **415**, 451–454 (2002).
7. Doebley, J., Stec, A. & Hubbard, L. *Nature* **386**, 485–488 (1997).
8. Langdale, J. A., Zelitch, I., Miller, E. & Nelson, T. *EMBO J.* **7**, 3643–3651 (1988).
9. Hall, L. N., Rossini, L., Cribb, L. & Langdale, J. A. *Plant Cell* **10**, 925–936 (1998).
10. Rossini, L., Cribb, L., Martin, D. J. & Langdale, J. A. *Plant Cell* **13**, 1231–1244 (2001).
11. Ueno, O. *Planta* **199**, 382–393 (1996).
12. Ueno, O. *Planta* **199**, 394–403 (1996).
13. Magnin, N., Cooley, B. A., Reiskind, J. B. & Bowes, G. *Plant Physiol.* **115**, 1681–1689 (1997).
14. Voznesenskaya, E. V., Franceschi, V. R., Kierats, O., Freitag, H. & Edwards, G. E. *Nature* **414**, 543–546 (2001).

Our virtual planet

Japan's Earth Simulator supercomputer could provide the most accurate models yet of the planet's climate and geophysics — but there are obstacles to realizing that potential. Robert Triendl reports.

On the outskirts of Yokohama, just outside Tokyo, a very special computer has been assembled. Housed in a building the size of an aircraft hangar, the Earth Simulator boasts some impressive statistics. Some 220 tonnes of cables link its 5,120 processors. Its combined memory is tens of billions times larger than that of a new desktop machine. And when running at full steam, it should be able to perform 40 trillion calculations per second, making it the most powerful computer in the world.

The Earth Simulator has been designed to model the climate and geophysics of the Earth with unprecedented accuracy. Initial tests of the machine began earlier this year, but serious questions about its operation remain unresolved. Getting the most out of such a machine requires new computer software and improved theoretical climate models, and Japan lacks the appropriate expertise. Collaborations could provide a solution, but the project's backers have yet to decide on how much access foreign researchers should be given. And at a price of about US\$370 million — plus up to \$50 million in annual maintenance costs — some observers suggest that the machine is costing the Earth, rather than simulating it.

The project has its roots in politics as much as in science. The machine uses vector processors, the manufacture of which is dominated by Japanese companies. These chips, which simultaneously perform the same mathematical operation on arrays of data, or vectors, were the mainstay of scientific computing at the beginning of the 1990s.

But the vector computer industry did not fare well during that decade. The performance

of the conventional processors used in desktop machines increased rapidly throughout the 1990s. Although each can only perform calculations on one data point at a time, hundreds or thousands of these relatively cheap chips can be wired up to create powerful 'parallel' supercomputers. In 1997, the import of Japanese vector machines into the United States was banned following a trade dispute between the two countries. So researchers in the United States — the biggest single market for supercomputers — turned increasingly to parallel systems.

Climate of change

By the late 1990s, Japan was looking for ways to help its struggling vector supercomputer manufacturers. A vector-based climate machine was an obvious choice. Climate researchers had not willingly abandoned vector machines, as it is difficult to adapt some climate models so that they run efficiently on parallel systems. And with a relatively weak base in climate science, the

Japanese government was aware that Japan was contributing little to the global effort to understand climate change. Plans for the Earth Simulator were launched in 1997, the year Japan hosted the climate conference that launched the Kyoto protocol on limiting greenhouse-gas emissions. The Japanese information technology company NEC, based in Tokyo, was given the contract to build the simulator.

By that time, large parallel systems had become considerably faster than single-processor vector supercomputers. To match the speed of these new systems, the Earth Simulator's designers took a hybrid approach. The result is a parallel machine built with vector processors, and wired with extremely fast interconnections.

This speed brings various advantages. Some researchers, for example, simply want to perform multiple simulations with existing climate models. Some outcomes of these models, such as the size of the temperature increase that greenhouse gases are expected to cause, are very sensitive to the details of the model. Running slightly different versions of the same model gives researchers an idea of the range of possible increases. A similar process is routinely used in weather prediction, says Thomas Stocker, a climate researcher at the University of Bern in Switzerland: "It must be done for climate predictions as well."

Another advantage is the possibility of higher-resolution models. Climate models divide the atmosphere into cuboid cells. The smaller the cell, the more realistic the modelling process can be, but smaller cells need more time to process. Time constraints



Bright sparks: simulating storms using the new machine will require refined convection models.

mean that existing climate supercomputers use cells with sides about 100 kilometres long when running global models. The Earth Simulator, in contrast, should be able to run simulations using cells with 20-km sides.

Smaller cells should help to capture features that existing models miss. John Mitchell, a modeller at the UK Meteorological Office's Hadley Centre for Climate Prediction and Research in Bracknell, west of London, says that features such as squall lines — rows of storms that form in the tropics — could be represented by the Earth Simulator.

The same holds for ocean simulations, where important currents in the Atlantic and at the Equator are not represented well in today's models. "Higher resolution requires increased effort," says Stocker. "But these investments are fully justified by the increase in the realism."

But moving to higher resolutions can be difficult. Mitchell cites the example of modelling convection in thunderstorms. The



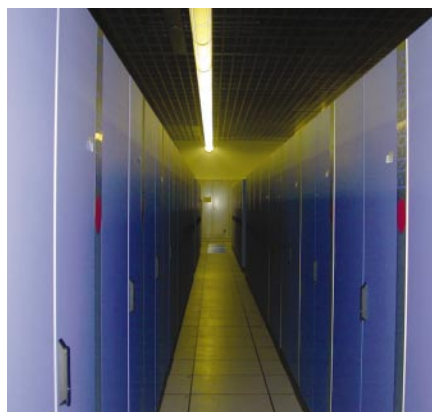
Tetsuya Sato: wants foreign researchers to get involved.

storms are driven by the heat released as warm, moist air rises and the water vapour condenses. Existing models simplify the process, as it occurs on a much smaller scale than the cells used in the model. But there is no guarantee that the same simplification will work for cells with 20-km sides, which are close to the scale on which the convection occurs. Unless current models are adapted, says Mitchell, they could start to produce unrealistic results.

Large parallel systems pose other problems. Simulations only run efficiently when the computational load can be spread evenly around all the processors in the system. In the case of climate simulations, each processor handles a set number of cells. But if one chip takes longer than others to complete a step in the model, the rest of the system must stop and wait. Even small delays can have dramatic effects on the system's performance, and such effects are impossible to eliminate completely.

Computer scientists in the United States have accumulated considerable experience in developing the software needed to minimize such problems, but the relevant expertise is rare in Japan. "With all the vector systems around, there aren't that many people in Japan who have experienced the sheer difficulty associated with large parallel computers," says Yasumasa Kanada, director of the University of Tokyo's computer centre.

Collaborations with foreign groups could help, but the Japan Marine Science and Technology Center in Yokosuka, close to Tokyo, which manages the Earth Simulator, has yet to draft guidelines for the machine's



Spaghetti junction: hundreds of tonnes of cables link thousands of vector processors to make the Earth Simulator the most powerful supercomputer in the world.

use. The Japanese government has only said that any work done on the system has to address the country's policy needs. This could, for example, mean focusing on climate changes specific to Asia, or the risk of earthquakes in and around Japan.

Officials at the Japan Atomic Energy Research Institute in Tokyo, the research agency that made the biggest contribution to the simulator's set-up costs, originally suggested that use be restricted to scientists based at Japanese institutes. But theoretical physicist Tetsuya Sato, the new director of the Earth Simulator Center, has been lobbying for considerable access to be given to foreign teams. Sato also suggests that time could be allotted to simulations in other disciplines, such as nuclear physics, nanotechnology and life sciences. With the system still undergoing



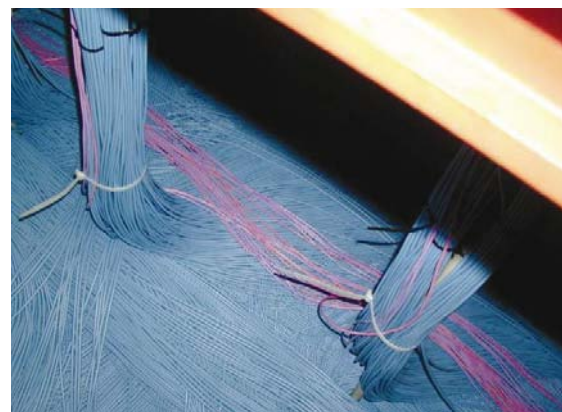
Keep it simple: Didier Paillard works with cheaper machines.

testing, and some of the software perhaps years away from completion, the debate could continue for some time.

Whether or not they are granted access to the Earth Simulator, climate researchers in Europe and the United States are already debating whether to lobby their paymasters for similar machines. Although most agree that fast machines are useful, some are asking whether their limited budgets would be better spent in other ways.

Elementary interaction

Didier Paillard, a climate modeller at the French Atomic Energy Commission's research centre in Saclay, near Paris, says that very simple models, which can be run on relatively cheap computers, can provide results different from, but just as useful as, those produced by high-resolution simulations. Rather than improving the resolution of the oceanic and atmospheric components that make up the models, Paillard argues that attention should be focused on



understanding the interactions between the two systems.

In the simple models favoured by Paillard, both the ocean and the atmosphere are represented by just a few cells. Although they cannot produce accurate climate simulations, these models offer insights into ocean-atmosphere coupling that more complex models miss. They can also reveal mistaken assumptions. "Ten years ago, the ocean was still assumed to limit climate change," says Paillard. "Now it is seen as a possibly very active system. Increased computing power didn't contribute much to elucidate this point."

Detailed outlook

But others defend the importance of resolution. "I used to be sceptical of the blind pursuit of resolution," says Richard Rood, a climate modeller at NASA's Goddard Space Flight Center in Greenbelt, Maryland. "But it brings phenomena closer to the scale on which observations are made." This, Rood argues, will allow observations to have a greater impact on models. "This is a critical bridge that must be crossed."

Whatever decisions other nations make, financial constraints in Japan mean that the Earth Simulator is likely to mark the end of the country's tradition of big computer hardware projects. "We are now looking at what kind of computing performance scientists really need, rather than at what is technologically feasible," says Masashi Shinozaki, who is responsible for high-performance computing at the Ministry of Education, Culture, Sports, Science and Technology.

The Earth Simulator is, nevertheless, a remarkable achievement. Regardless of the politics behind its inception, the result is a machine of unprecedented capability. And if Sato gets his way, its impact could be felt beyond the limits of nationality and disciplines that some of its original backers proposed. It is not too late for the Earth Simulator to have a truly global impact. ■

Robert Triendl is a freelance writer, and works for Asia Technology Consulting in Tokyo.

♦ www.es.jamstec.go.jp

How affluence could be good for the environment

Intensive farming means the United States uses less land to feed more people than ever.

Sir — A current issue in assessing the environmental effects of economic growth is the per capita use of land in developed versus developing countries. It is often stated that countries with higher incomes demand much more land per person to sustain their higher consumption rates. E. O. Wilson¹, for example, cites the Worldwide Fund for Nature (WWF)² in asserting that the “ecological footprint” of each person in the United States is 9.6 ha, against a global average of 2.1. From this, Wilson concludes it would require “four more planet Earths” for everyone to live as Americans do.

The accurate, long-running land-use statistics compiled by the US Department of Agriculture Economic Research Service³ offer a way to check these assumptions. Using these data, I have calculated per capita US land requirements over time.

Food production represents by far the greatest impact of humans on land worldwide⁴: about 10% of the world's 13 billion ha of land is used to grow crops, 25% for permanent pasture. The continental United States has a land area of 2.8 ha per person at present. Crops are grown on 0.52 ha per capita; pasture, grazed forest land and grazed arable land add an additional 1.17 ha per capita. By comparison, forested lands, parks and wildlife areas total 0.75 ha per capita.

What is most striking is that the total area used for US food production has remained virtually unchanged since 1945. As a result, from 1945 to 1997 per capita cropland use declined by 50%, even as population grew by 90% and real per capita disposable personal income grew by 177%.

The United States is a substantial net exporter of food. For all high-income countries taken together, average cropland area is 0.4 ha per capita. By comparison, sub-Saharan Africa, which has the world's lowest nutritional status, averages 0.3 ha per capita of cropland and 1.7 ha per capita of grazing land⁵.

Thus, developed nations provide more calories per person, and very substantially more meat and dairy products per person (per capita meat consumption in the United States is 10 times that in sub-Saharan Africa), while using only slightly more cropland per capita than the poorest countries and less total agricultural land (arable plus grazing) per capita.

The reasons for this are not surprising. Farmers in poorer countries attain much lower yields per ha, largely as a result of very low fertilizer use: 14 kg of fertilizer per ha of cropland in sub-Saharan Africa

versus 114 kg per ha in the United States is typical⁵. As countries become wealthier, yields increase substantially. Since 1945, US yields per ha of wheat have more than doubled and of maize have more than tripled (ref. 4); the world average yield per ha of maize is now half the US yield. Use of chemical fertilizer and of hybrid and other high-yielding varieties of grains could let developing countries match Western diets with little or no increase in land use.

The growth in urban areas and other uses of land that come with growing affluence add an insignificant amount to land requirements. In the United States, these uses grew from 0.11 ha per capita in 1945 to only 0.13 ha per capita in 1997.

The WWF derived its very high per-capita land requirements for developed countries by including in its “ecological footprint” the theoretical amount of forested land that would have to be added to the Earth's surface to absorb CO₂ emissions from those countries. But this is an apples-and-oranges comparison; planting forest is but one of many ways that CO₂ emissions could be reduced (substituting nuclear energy, for example,

adds zero land requirements), and in no way constitutes a limit on available land resources of the kind Wilson infers.

In the larger picture, it is worth noting that economic growth in developed countries has become increasingly decoupled from dependence on natural resources. Farm production was 7.2% of US gross domestic product in 1945 but only 1.1% in 1997. Paul Ehrlich⁶ has argued that environmental impact is proportional to population times affluence. But as far as land requirements are concerned, it appears to be proportional to population divided by affluence.

Stephen Budiansky

Black Sheep Farm, 14605 Chapel Lane, Leesburg, Virginia 20176, USA

1. Wilson, E. O., *The Future of Life* (Knopf, New York, 2002).
2. Worldwide Fund for Nature, *Living Planet Report 2000* (Gland, Switzerland, 2000).
3. US Department of Agriculture, *Agricultural Resources and Environmental Indicators, 2000*. USDA Economic Research Service, Resource Economics Division (Washington, 2001).
4. Waggoner, P. E., “How much land can ten billion people spare for nature?” Council for Agricultural Science and Technology, Task Force Report No. 121 (Washington, 1994).
5. United Nations Food and Agriculture Organization, FAOSTAT Database; apps.fao.org.
6. Ehrlich, P. R. & Ehrlich, A. H. *The Population Explosion* (Simon & Schuster, New York, 1990).

Training to think globally

Sir — Of the approximately 300 molecular-biology PhD candidates at Harvard Medical School, only two carry out research on malaria, a disease that kills one child every 40 seconds (see J. Sachs and P. Malaney, *Nature* **415**, 680–685; 2002). An often-cited reason for the lack of interest in pursuing research into diseases prevalent in developing countries is the limited amount of funding allocated to this type of project (with the exception of HIV/AIDS).

However, a more significant factor responsible for this trend could be the nearly complete absence of any discussion related to diseases common in developing countries during the course of our graduate training. For example, we escaped our first year of graduate school without any exposure to current malaria or tuberculosis research.

If those responsible for our training choose to focus almost entirely on diseases most common in the North, such as heart disease or obesity, then the next generation of scientists they train are likely to do the same. Making more effort to teach molecular-biology PhD candidates about the mechanisms of tuberculosis infection or strategies for malarial vaccine design will cost nothing and will go a long way

towards creating a new generation of globally minded scientists.

Franklin Huang*, Matt Stremlau†

**Washington University School of Medicine, 660 S. Euclid Avenue, St Louis, Missouri 63110, USA*

†Harvard Medical School and Dana Farber Cancer Institute, Department of Cancer Immunology and AIDS, Jimmy Fund Building 617, 44 Binney Street, Boston, Massachusetts 02115, USA

Hamilton built on work by Haldane and Fisher

Sir — Olivia Judson's perceptive review of W. D. Hamilton's second volume of Collected Papers (*Nature* **416**, 17–18; 2002) opens with a well-known misapprehension that should be allowed to die. The “discovery of inclusive fitness” was not his, but R. A. Fisher's and J. B. S. Haldane's, to whom Hamilton referred in his papers. It is in both their books, Fisher's *The Genetical Theory of Natural Selection* (Clarendon, Oxford, 1930) and Haldane's *The Causes of Evolution* (Longman, London, 1932), as well as their earlier papers and essays.

A. W. F. Edwards

Gonville and Caius College, Cambridge University, Cambridge CB2 1TA, UK

Thoughts of a rationalist

Provocative ideas about science from a reductionist viewpoint.

Facing Up: Science and Its Cultural Adversaries

by Steven Weinberg
Harvard University Press: 2001. 304 pp.
£17.95, \$26

John Polkinghorne

Collections of occasional writings do not always make good reading. In fact, they often seem little more than a lazy person's way of compiling a book. Steven Weinberg's latest publication, consisting of miscellaneous non-specialist pieces culled from his output over the past 17 years, fares a great deal better than most such offerings. There is certainly an unwelcome degree of overlap and repetition between the essays, but Weinberg writes well and his clear style and strong opinions hold the reader's attention. He says he has "a taste for controversy", and this punchy and provocative writing certainly bears this out.

Weinberg describes himself as "rationalist, reductionist, realist and devoutly secular", using all these adjectives in a very strong sense. I can only share one of these qualities with him: realist. Weinberg has been a vigorous participant in the so-called 'science wars', attacking strong accounts of the social formation of science and defending the claim that science achieves a verisimilitudinous understanding of what the physical world is actually like. He frequently asserts the equal reality of quarks and rocks. Weinberg acknowledges that reality is a hard concept to define, but he regards that as simply a technical problem for philosophers to solve, which should not give pause to scientists. I agree with him, as I do with his oft-repeated criticisms of the stance taken by Thomas Kuhn in these matters.

When it comes to reductionism, Weinberg



takes a very hard line. He does not want to be a fundamental physics imperialist, in the sense of claiming that the practice of other disciplines could actually be reduced to corollaries of particle physics, but he does see his subject as providing an in-principle explanatory basis that needs no further augmentation. Weinberg calls this view "Newton's dream". While everyone in the scientific world would like to think that they have Sir Isaac on their side, this assignment of Newton to the hard reductionist position seems a highly curious judgement on a man who spent so much time on alchemy and

the interpretation of prophecy, and who was a deeply religious person who believed in the causal powers of spirits (including, as he supposed, occasional corrections to the motions of the Solar System).

The acid test of reductionism is how it treats consciousness. Weinberg is very critical of ideas of emergence such as those defended by Philip Anderson with his well-known epigram 'More is different'. (Emergentists hold that a system will, as it grows more complex, display properties, such as consciousness, that cannot be explained in terms of the behaviour of its components.) Weinberg acknowledges that consciousness is very far from being understood, but so, he says, is the weather. Here he fails to recognize an important distinction between two kinds of emergence: the comparatively unproblematic (such as weather systems) and the problematic (such as mental systems). Weather systems involve exchanges of energy, and there is no insuperable difficulty in believing — to use a favourite Weinberg metaphor — that here the 'arrows of explanation' point down ultimately to basic physics. On the other hand, mental experiences — such as seeing red — appear to be qualitatively different from exchanges between neurons, let alone interactions between quarks.

Strong reductionism of the Weinberg kind is, in fact, a metaphysical assertion. If it

New Journals

This year, *Nature's* annual new journals review supplement will appear in the 7 November issue. Publishers and learned societies are invited to submit journals for review, as well as details of any eligible electronic journals, taking note of the following criteria:

- Journals must have first appeared during or after June 2000 and published at least four separate issues by the end of June 2002.
- Journals covering any aspect of science are eligible, although those dealing with clinical medicine and pure mathematics are excluded, as are newsletters and publications of abstracts.
- Frequency of publication must be at least three times a year.
- The main language must be English.
- Deadline for submissions is 15 July 2002.

For each eligible title, please send at least four different issues (the first, the most recent and any two others), together with full details of subscription rates, to Mary Purton, *Nature*, The Macmillan Building, 4 Crinan Street, London N1 9XW, UK. Tel: +44 (0)20 7843 4567. Fax: +44 (0)20 7843 4596/4763. E-mail: m.purton@nature.com

Flights of fancy

The 980 species of bat have many remarkable features and show huge diversity. It is well known that they use echolocation to navigate and detect prey — hence the large ears of the long-eared bat, *Plecotus* sp., shown here. But the role of bats in seed dispersal and controlling insect populations is often overlooked, and without pollination by bats there would be no tequila. Now they are under threat, from habitat destruction and the use of pesticides, a recurring theme in *Bats* (The Natural History Museum, £9.95) by Phil Richardson, who was recently appointed by Britain's National Trust as their first bat conservation officer.

is to be defended, this has to be on metaphysical grounds that go beyond the simply scientific. Although the judgement may seem odious, to imply otherwise is indeed an act of physics imperialism.

Another significant metaphysical issue is the status of values, such as ethical insights. Weinberg frequently affirms the importance of values, and he does not claim that they are something for which science itself can establish a basis. He sees them as human creations, by which we make for ourselves a little island of meaning in the hostile and meaningless Universe that surrounds us. There is a certain nobility in this bleak view, but I believe it to be mistaken. Our ethical intuition that abusing children is wrong is not simply something we have decided to assert. It is an actual fact about the complex reality within which we live, a form of knowledge whose origin is a question that requires very careful consideration. A true rationality has to take the reality of this kind of personal experience as seriously as it does science's impersonal encounter with the world.

This latter kind of reasoning is part of my motivation for being a religious believer. Weinberg is well known as an implacable opponent of religion. One of the essays in this volume arose from a conference to which we both contributed, which was convened to discuss whether there are any signs of inbuilt design in the Universe. Much of the argument related to the 'Anthropic Principle', the surprising recognition that a universe capable of generating carbon-based life has to satisfy very tight conditions on the strengths of the forces of nature operating within it. Weinberg refers to these matters from time to time, particularly noting the astonishingly tiny value (10^{-120} of its naturally expected strength) that the cosmological constant has to take in an anthropic world. He takes the multiverse way out, avoiding any whiff of theism by appealing to the highly speculative and metaphysically prodigal notion that our world is just, by chance, the anthropically

fruitful member of a vast portfolio of diverse universes.

When Weinberg comments more generally on religion, the tone is highly polemical. He is right to draw attention to the horrors of religious wars and persecutions, for which a believer must feel shame and penitence, even if such terrible happenings are by no means the monopoly of religions. Yet that dark side is far from being the whole story, for much good and achievement (including initially modern science, some would claim) have also come out of religious belief. Weinberg rightly raises the issue of suffering, asking "how does free will account for cancer?" It does not, but an evolving world (theologically, a creation allowed to 'make itself') is one in which cellular mutations necessarily bring about not only new forms of life, but also malignancy.

This is an interesting book, with a lot to question and (I believe) disagree with, but is well worth reading. ■

John Polkinghorne is the retired president of Queens' College, Cambridge, CB3 9ET, UK. His latest book, published this month, is *The God of Hope and the End of the World* (Yale University Press, 2002). He was awarded the Templeton Prize for Progress in Religion on 14 March 2002.

It's not all in the genes

Phenotypic Plasticity: Beyond Nature and Nurture

by Massimo Pigliucci

Johns Hopkins University Press: 2001. 384 pp. £25, \$75

Ralph Tollrian

Phenotypic plasticity describes the ability of a genotype to form different phenotypes in distinct environmental conditions. Although originating in the field of evolutionary biology, phenotypic plasticity is a truly interdisciplinary topic, lying at the

intersection between most of today's important biological disciplines, from evolution to molecular biology and, of course, ecology. Phenotypic plasticity occurs in many traits ranging from morphology through developmental biology to physiology and behaviour, and can be found in all classes of organism. Nearly every biological discipline deals with some aspects of phenotypic plasticity.

Massimo Pigliucci's *Phenotypic Plasticity: Beyond Nature and Nurture* is the first volume in a new series entitled "Syntheses in Ecology and Evolution" edited by Sam Scheiner. Both scientists are leading experts in the field of phenotypic plasticity, and their names are a guarantee of a high standard.

It is a good time for a book on this topic. Following the renaissance of the 'reaction norm' (a term coined by Woltereck in 1909, referring to a set of phenotypes expressed by an individual genotype in response to an environmental gradient), a wealth of papers in the past two decades has described phenotypic plasticity and genotype-environment interactions. Whether phenotypic plasticity is a target or a by-product of selection has been hotly debated, which has led to the realization that it can be both. Now, thanks to the rapid development of molecular techniques, the field is at a stage where new questions can be asked. What are the developmental mechanisms? What are the pathways? How does a genotype map into a phenotype? When does selection lead to the evolution of distinct morphs or even species, and when to phenotypic plasticity? Pigliucci pinpoints these key questions with a focus on mechanisms and processes.



Working model: thale cress (*Arabidopsis thaliana*) is often used in studies of phenotypic plasticity.

A strength of the book is Pigliucci's attempt to outline avenues for future research. Following a general introduction to phenotypic plasticity, he gives a broad overview of the recent contributions of genetics, molecular biology, developmental biology, ecology, behavioural biology, evolutionary biology and theory to a modern synthesis of phenotypic plasticity. He then emphasizes, using *Arabidopsis thaliana* as an example, how research on phenotypic plasticity can lead to insights into other fields of evolutionary biology. In the final chapter, he extends the scope of the book beyond 'nature and nurture' to socio-political aspects of the question of how genes and environment influence human intelligence.

Phenotypic plasticity has been a battlefield for many, sometimes semantic, controversies. Pigliucci gives a balanced presentation, separating his own from alternative opinions. Nevertheless, there remain many

anchor points for stimulating discussions. For example, he emphasizes that plasticity in morphology, physiology and behaviour share a general framework, but suggests that one should distinguish between these fields, because timing and mechanisms differ. An alternative view would be that it is exactly the interplay between these differences that creates a promising area of research.

The book demonstrates Pigliucci's unusually broad knowledge of the topic. However, it is impossible to cover all aspects of phenotypic plasticity within the publisher's page limits for a single book. Consequently, he deals with some topics only briefly. For example, the factors that favour the evolution of phenotypic plasticity (R. Tollrian and C. D. Harvell, *The Ecology and Evolution of Inducible Defenses*, Princeton University Press, 1999) would, in my opinion, have benefited from more space.

Although there is some overlap with

an earlier book on phenotypes by C. D. Schlichting and Pigliucci (*Phenotypic Evolution: A Reaction Norm Perspective*, Sinauer, 1998), *Phenotypic Plasticity* is the most comprehensive book on this topic. It provides both a solid basis for understanding the subject and an inspiring synthesis of the current state of the discipline, and so can be equally recommended for students starting their research, for experts in the field and for all scientists generally interested in phenotypic plasticity. Hardly anyone will read this book without gaining new insights or new inspirations. The book is a 'must read' in the fields of evolution and ecology, and as such is an ideal topic for seminars. I highly recommend it and look forward to the next volumes in Scheiner's series. ■

Ralph Tollrian is in the Department of Evolutionary Ecology, Ludwig-Maximilians-Universität München, Karlstrasse 23-25, München D-80333, Germany.

Science in culture

Infinite ideas

A theatrical contemplation of infinity makes full use of industrial space.

Roald Hoffmann and Sylvie Coyaud

There's science in the theatre — Michael Frayn's *Copenhagen*, Peter Parnell's *Q.E.D.*, and *Oxygen* by Carl Djerassi and one of us. And now there is Luca Ronconi and John Barrow's *Infinities*. Italy's most innovative director has staged in Milan a remarkable teatro-architectonic meditation on, indeed, infinity.

Five pieces of text were written for the play by John Barrow, quoting Jorge Luis Borges, the physicist–novelist Alan Lightman and the mathematician Georg Cantor, among others. The texts are integrated by Ronconi's vision into a journey of engrossing variety. In Milan's industrial section, Bovisa, La Scala Opera company had refurbished some old buildings for its workshops. This is Ronconi's setting. Or better said, settings, for the audience moves from one space in the vast complex to another.

One begins in a bar atrium, rising eight floors: L'Albergo Infinito, or the Hilbert Hotel as mathematicians may know it, whose manager must accommodate an infinite number of guests. He rehearses different solutions and proves some theorems along the way. Next, in a windowless room, women young and old alternate between desire and dread of living forever. Elsewhere, the paradoxes of time travel are paraded around a large hangar, that has a train carriage suspended in mid-air at one end and crates of books at the other.

The actors pass through the audience, and in one effective scene, where Cantor's work and life is told, we sit at tables with students/actors from Milan Polytechnic University. Other actors walk on the tables, climb ladders, or speak



suspended upside-down from a conveyor belt. Their tight face masks, black and white costumes and moving platforms are all Ronconi's hallmark.

The most successful, indeed unforgettable, of the scenes accomplishes what we thought could not be done — it recreates on stage Borges' Library of Babel. A cavernous space, several stories high, is filled with cubicles and empty bookcases, two slanted mirrors stretching them effectively to infinity. Narrators appear on passages and bridges, making a connection not

only to Borges' dismal library, but also to Andrea Palladio's theatre at Vicenza.

The audience is admitted in groups of 60 to 80 every 15 minutes, and moves through the five sets in roughly two hours. Meanwhile the actors, a large cast, also rotate, adding to the sense of endless motion.

Do we learn any mathematics? Yes. Does it matter? Not at all. This is a theatre of space and time, and of ideas. It is abstract, for the dramatic moments come not in confrontation of human beings, in that silence between words and action where a gesture stretches time; rather, they come from the tug-of-war between hope and despair, or comic failure. To the mind bold enough to pursue it, we are shown, infinity promises unlimited freedom — and delivers madness.

The Piccolo Teatro production, sponsored by Fondazione Sigma Tau, ran for just three weeks in Milan, ending on 28 March. On 19 April it will inaugurate the "Ciutat de les Arts Escéniques", created by Spain's leading architect, Santiago Calatrava, in another former factory, the Nave de Sagunto in Valencia. Given the integration of theatre with setting, the play cannot be the same. One is curious to see how the Spanish director Vicente Genovès will translate Ronconi's vision into the soaring Nave, and whether the cast, coached by the Greek actress Irene Papas, will match the bravura of the Piccolo's performance. ■

Roald Hoffmann is professor of chemistry at Cornell University, Ithaca, New York. Sylvie Coyaud is a science journalist based in Milan.

For details on the production, in Italian, see <http://www.piccoloteatro.org/infinities>

The art of chemistry

Stephen J. Lippard

In many respects, synthesis is the heart of chemistry. The creativity involved in combining the elements of the periodic table into unprecedented molecules and materials distinguishes chemists from other scientists. But wielding the chemical paintbrush in a controlled manner is a challenge.

Consider the synthesis of natural products. In assembling a complex organic molecule such as an antiviral or anticancer agent, the synthetic chemist must plan and execute multiple reactions, typically in a stepwise manner, until the desired compound is obtained. Success often requires the ability to react a single part of a molecule to the exclusion of others. Such 'chemoselectivity' is typically achieved with a protecting group that can later be removed without destroying any of the product. The labour-intensity of this process is analogous to being dressed in medieval armour and having to wash with a fire hose, cleaning the face by raising the visor, the hands by removing the gauntlets, and so on.

A key goal in synthetic organic chemistry is the design of reagents to achieve chemical transformations at specific sites in a molecule without protection and deprotection steps. To achieve one stereoisomer to the exclusion of its mirror image is a further, often more difficult, challenge, as is the development of catalysts that promote an efficient chemical transformation.

The synthetic inorganic chemist faces a different but equally formidable task of control. Consider the goal of preparing a catalyst for a chemical transformation that occurs in living organisms at ambient temperature and pressure, such as the conversion of

atmospheric nitrogen to ammonia, natural gas to methanol, or water to hydrogen. The first of these catalysts would significantly assist in the fertilization of crops; the second would allow methane to be transported much more economically from remote gas wells to urban areas; and the third would provide a key ingredient for fuel cells from a non-carbon, replenishable fuel source.

The enzymes that catalyse such reactions in nature typically operate at kilohertz frequencies: many biological chemical reactions convert substrates to products at rates of roughly 1,000 per second. Almost half of these enzymes require metal ions to achieve their catalytic functions — usually Mg, Ca, Mn, Fe, Co, Ni, Cu and Zn — typically embedded in a three-dimensional protein matrix that ensures tightly controlled alignment of the reactants to achieve the desired formation and release of the product molecule(s). Outside the protein milieu, these metals in their typical +2 and +3 oxidation states form complexes that exchange their ligands rapidly — they are kinetically labile.

These ions seem, therefore, to have been adopted by living organisms precisely because they can bind to substrates and release products on a timescale commensurate with biological requirements. Notably absent from the list are second- and third-row transition metals, such as Ru, Pd and Pt. Complexes containing these metal ions are important as industrial catalysts, but their rates of ligand exchange are much lower than those of their first-row counterparts. With a few exceptions, such as Mo and W, such elements have therefore not been selected for biological functions.

A great challenge for the synthetic inorganic chemist who is attempting to mimic an enzymatic transformation is to create an environment that stabilizes the ligand composition at the metal ion, while offering labile sites for catalysis. In a biological molecule, ligands are typically supplied by amino-acid side chains and are therefore held in relatively fixed positions. Disassembly or rearrangement to thermodynamically more stable entities is kinetically prohibited by energy supplied by the folding of the protein into its specific three-dimensional shape. In creating a functional synthetic analogue of the protein's active site, however, the chemist is faced with the daunting task of overcoming the tendency of the desired target molecule to rearrange or react with components of the solution to form a thermodynamically stable, undesired side product with no functional capability. The system is thus under thermodynamic rather than kinetic control, because of the labile nature of the metal ions involved.

Yet in organic chemistry, one more

Chemical synthesis

Success in the synthesis of natural products often requires the ability to react a single part of a molecule to the exclusion of others.

frequently has kinetic control of the molecule. For example, in preparing a catalyst with a phenyl (C_6H_5) ring as the framework on which a reactive moiety is appended, the organic chemist does not have to worry about undesired transformations involving the spontaneous cleavage of C–C and C–H bonds in the phenyl unit, because these bonds are kinetically stable. The use of designed peptides is one approach to the problem of creating new molecules as ligands for transition-metal ions that provide parallel kinetic stability for inorganic catalysts but allow for functional sites.

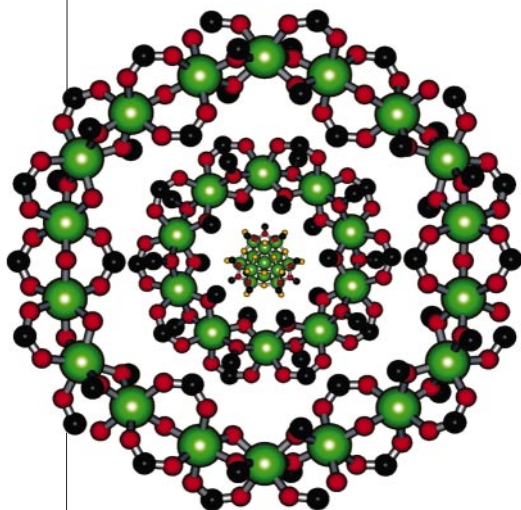
On occasion, the inability to control a desired chemical transformation gives rise to new molecules that can be objects of aesthetic beauty as well as potential usefulness. One example is the serendipitous synthesis of carboxylate-bridged polyiron compounds (see figure), including the 'molecular ferric wheel', the 'molecular 18-wheeler' and an open-shell Keggin ion. I have not seriously pursued the rational design of such molecules because I do not understand the principles of control involved. In cases in which many interactions are possible, particularly many non-covalent, subtle interactions working in concert, the chemical paintbrush can seem to have a mind of its own. It is often impossible to predict the factors that will dominate and dictate the nature of the product.

With the use of kinetically more inert transition-metal ions in conjunction with bridging ligands, however, it is possible to assemble molecular squares, cubes and other polygons and polyhedra, as well as porous solids with potential for a range of applications. Possible examples include new materials for cleaning waste water, smart detectors for chemical and biological warfare agents, and components for molecular electronics devices.

Stephen J. Lippard is in the Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

FURTHER READING

Lippard, S. J. *Chem. Eng. News* **78**, 64–65 (2000).
DeGrado, W. F., Summa, C. M., Pavone, V., Nastri, F. & Lombardi, A. *Annu. Rev. Biochem.* **68**, 779–819 (1999).
Taft, K. L. *et al.* *J. Am. Chem. Soc.* **116**, 823–832 (1994).
Watton, S. P. *et al.* *Angew. Chem.* **36**, 2774–2776 (1997).
Leininger, S. *et al.* *Chem. Rev.* **100**, 853–907 (2000).



Chemical revolution: (from inside) the open-shell Keggin ion, ferric wheel and molecular 18-wheeler.

Clearing a path for nerve growth

Lars Olson

After spinal-cord injury, severed nerve fibres face a thicket of obstacles as they try to regenerate. Researchers have used a bacterial enzyme to help prune these obstacles in rats.

Spinal-cord damage is tragically common, affecting three to five people out of every 100,000 in the United States and similar numbers in other countries. The consequences for both the injured person and their families and friends can be devastating. Injuries typically occur in young adults, and of the 11,000 or so new cases each year in the United States, 80% are men who were involved in driving accidents, violence or falls. Most victims survive, but are completely or partially paralysed for the rest of their lives.

It has long been thought that such traumatic injuries are incurable, but intense experimental work is beginning to suggest that they might one day be treatable at least. In the latest advance, Bradbury and colleagues¹ provide evidence (page 636 of this issue) that infusion of chondroitinase ABC — a bacterial enzyme that trims the carbohydrate side chains off large extracellular proteins — enables severed nerve fibres to regenerate after the spinal cord has been crushed in rats. The authors used forceps to squeeze the nerve fibres in the so-called dorsal columns of the spinal cord until nerve fibres carrying sensory information from the limbs to the central nervous system, and fibres conveying motor (movement) information to the limbs, had been cut. This severely impairs limb function. In animals that were then treated with chondroitinase ABC, there was convincing evidence of nerve-fibre regeneration at the cut nerve stumps. Although the growth of sensory nerve fibres was not enough to lead to a clear recovery of sensation, the authors have electrophysiological and behavioural evidence of remarkable motor recovery.

Imagine that you are as small as the tip of a nerve fibre, a few thousandths of a millimetre in diameter, and have to advance in the narrow regions between cells — the extracellular space — in a recently damaged tissue. You would find that these spaces are by no means open territory, but rather like an endless overgrown shrubbery, consisting of many different long, and sometimes branched, molecules that form an intertwined network, the extracellular matrix (Fig. 1). Some of the molecules are anchored to cells, some hold adjacent cells together, and many are simply deposited outside the cells. The terrain becomes particularly hostile after injury because new classes of cells

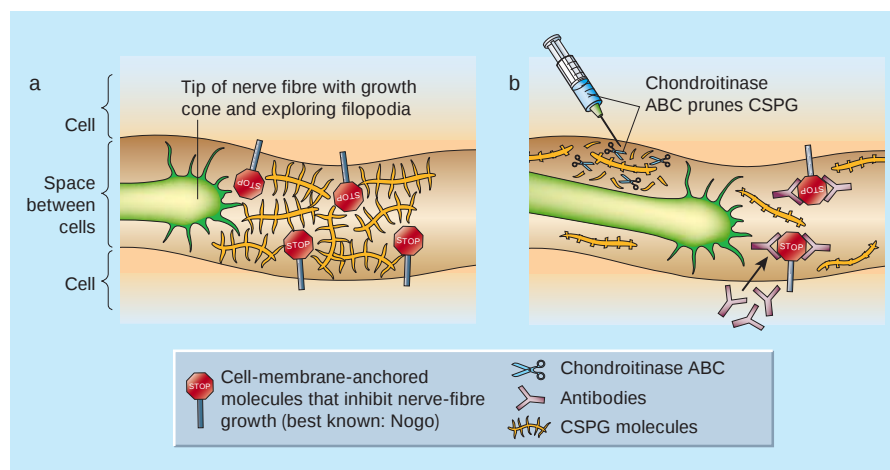


Figure 1 Obstacles to nerve regeneration. **a**, A severed nerve is shown on the left, in the extracellular space between cells in a wounded spinal cord. The nerve end has just formed a 'growth cone', which is exploring the territory by constantly forming and retracting minute processes (filopodia). The extracellular space contains two obstacles to nerve growth: chondroitin sulphate proteoglycan (CSPG) molecules, which have a backbone and many side branches that block the way; and special cell-membrane-anchored proteins that actively stop growing nerve fibres. The best known of these is Nogo. **b**, Bradbury *et al.*¹ show that the bacterial enzyme chondroitinase ABC can prune the side chains of CSPGs, clearing the way for growing nerve fibres. It has previously been shown that Nogo can be blocked by specific antibodies. The antibodies and enzyme could potentially be used in combination, along with proteins that stimulate nerve growth.

arrive, and a scar that is burdened by increasing deposition of matrix molecules builds up^{2,3}. To advance, you would need a machete.

The chemistry of the extracellular matrix is complex, but one key class of molecules are the chondroitin sulphate proteoglycans (CSPGs). These are made up of a protein core — the chondroitin sulphate — that is equipped with many side chains consisting of carbohydrate building-blocks such as glycosaminoglycans. Bacteria such as *Proteus vulgaris* have evolved enzymes, including chondroitinase ABC, that can prune CSPGs by removing glycosaminoglycans. Perhaps this helps the bacteria to invade their animal hosts.

Knowing all this, Bradbury *et al.*¹ wondered whether chondroitinase ABC could be used as a molecular machete to clear the way for nerve regeneration in injured rat spinal cords. After crushing the dorsal columns of rats at the level of the fourth cervical vertebra, the authors repeatedly infused either active chondroitinase ABC or a control protein, delivered through a thin tube, to the site of injury. Then, using an antibody that binds to the CSPG core but not the intact molecule, they showed that chondroitinase ABC effec-

tively prunes CSPGs in the spinal cord (Fig. 1). CSPGs remained intact in controls.

In the treated rats, sensory neurons in the dorsal root ganglia corresponding to the fourth cervical vertebra upregulated a nerve-growth marker protein, GAP-43, and the cut extensions (axons) of these neurons in the spinal cord regenerated by up to 4 millimetres towards the brain¹. Similarly, the important dorsal corticospinal tract, completely severed by the crush, regenerated downwards from the site of injury in the spinal cord. This was paralleled by a return of electrical signalling between the brain's cerebral cortex and the spinal cord, albeit at a reduced strength and somewhat slower speed. Bradbury *et al.* also showed that if they cut the spinal cord again, stimulation of the cerebral cortex no longer elicited signals in the cord.

In attempting to repair experimentally damaged spinal cords, it is crucial to see how well the animals recover. Bradbury *et al.* went to great lengths to test this. They show that rats that had been treated with chondroitinase ABC recovered normal or near-normal walking behaviour in two tests — beam and grid walking. But sensorimotor functions — becoming aware

of and removing a piece of adhesive tape — were not significantly recovered.

The results¹ are promising nonetheless, and it might prove possible to integrate the use of chondroitinase ABC with other strategies (compare with ref. 4). Let's consider the worst-case scenario, in which the spinal cord has become completely divided, leaving a gap between two stumps. While scar tissue invades the gap, initial oedemas (swellings) in the stumps themselves may transform into growing, fluid-filled cysts⁵. Yet both stumps contain severed axons with an intrinsic ability to regenerate — if the conditions are right. The problem can be likened to opening up traffic on a blocked-off, overgrown and partly demolished road. We may need to remove scar tissue, empty cysts, build physical bridges such as nerve grafts across a void⁶, and perhaps provide cellular surfaces that prompt nerves to grow^{7–9}.

To clear the road, chondroitinase ABC could be delivered to prune back the extracellular-matrix shrubbery. Chemical stop signs — such as those composed of the nerve-growth-inhibitory protein Nogo, expressed by supportive glial cells — could be covered over by antibodies and thus neutralized¹⁰. Chemical gradients of nerve-growth-activating molecules may need to be established to stimulate and to direct growing nerve fibres. Finally, to protect newly regenerated nerve fibres and increase signalling speed, cells that provide fibres with a new insulating sheath (composed of a lipid, myelin) might be needed. It may prove easier to induce long axon growth in white matter than in grey matter. Grey matter is made up mainly of nerve-cell bodies, with some other cells, and the complex network of contacts between nerve-fibre branches. White matter consists largely of myelin-coated nerve fibres. Any fibres in grey matter tend to branch greatly, so the neuron's need for growth factors from underlying cells may be satisfied long before it reaches its target area. But fibres in white matter branch less, so can perhaps grow further before they are adequately supported.

What could go wrong by using chondroitinase ABC? Not much, it seems. Highly purified preparations have no protein-degrading activity and so would not destroy crucial proteins in the body. However, CSPGs come in many forms, and some might have specific roles in the central nervous system, for example to guide growing axons. In grey matter, many nerve-cell bodies are surrounded by nets of particular CSPGs, which might be important in neuronal contacts. These nets would degrade upon enzyme treatment, and take a long time to rebuild.

That said, treatment with chondroitinase ABC has also proved effective in treating brain injuries¹¹, and constitutes one hopeful avenue to promoting nerve regeneration.

Several other treatment strategies (see, for example, ref. 12) have also shown promise in different experimental models of spinal-cord injury. The bad news is that none of these techniques has led to complete recovery. Fortunately, however, many of them can be combined, perhaps in ways discussed above. In the short term, the prognosis for people with complete spinal-cord injury remains grim. Yet, looking further into the future, we can perhaps allow ourselves to be a bit more optimistic. ■

Lars Olson is in the Department of Neuroscience, Karolinska Institute, 171 77 Stockholm, Sweden.

Plant biology

The first harvest of crop genes

Michael Bevan

Draft sequences of the rice genome have been produced by two groups. The drafts will be an invaluable resource for research on the genomes of other plants, the cereals in particular.

Last week's publication in *Science*^{1,2} of two draft sequences of the rice genome is another milestone in the inexorable progress of genome sequencing. The work is significant on three counts. Rice (*Oryza sativa*) is the staple in the diet of nearly a third of the world's population; the sequences provide a reference for all the other cereals, which collectively account for an overwhelming proportion of human and domesticated animal feed; and this is the first representative of the monocotyledonous plant lineage to be sequenced. Together with the dicotyledonous plants, the monocots (grasses, palms, rushes, and so on) constitute the flowering plants, the most diverse and numerous of all plants.

One of the draft sequences has been produced by Syngenta (a Swiss agricultural biotech company)¹; the other comes from a collaboration between the Beijing Genomics Institute and other Chinese academies². Both reports are initial analyses of assembled whole-genome shotgun sequences, involving random sequencing of the entire genome and assembly of the sequence into contiguous regions, or contigs. The Syngenta group worked with the subspecies *japonica* Nipponbare (a popular rice cultivar in Japan), whereas the Chinese team sequenced a hybrid *indica* variety, the major subspecies grown in China and most other regions. Draft DNA sequence is generally accurate enough to identify genes and to estimate genome size. In this case, the sequences reveal a genome of 362–389 million base pairs (Mbp), and a conservative number of between 33,000 and 44,000 genes. Two other important studies of the rice genome provide independent evidence, based on the construction of a near-complete physical map³ and a map of expressed genes⁴, that the gene-

e-mail: lars.olson@neuro.ki.se

1. Bradbury, E. J. et al. *Nature* **416**, 636–640 (2002).
2. Lemons, M. L., Howland, D. R. & Anderson, D. K. *Exp. Neurol.* **160**, 51–65 (1999).
3. Davies, S. J. et al. *Nature* **390**, 680–683 (1997).
4. Olson, L. *Nature Med.* **3**, 1329–1335 (1997).
5. Josephson, A. et al. *Neurosurgery* **48**, 636–645 (2001).
6. Cheng, H., Cao, Y. & Olson, L. *Science* **273**, 510–513 (1996).
7. Li, Y., Field, P. M. & Raisman, G. *Science* **277**, 2000–2002 (1997).
8. Ramon-Cueto, A. et al. *Neuron* **25**, 425–435 (2000).
9. Hofstetter, C. P. et al. *Proc. Natl Acad. Sci. USA* **99**, 2199–2204 (2002).
10. Fouad, K., Dietz, V. & Schwab, M. E. *Brain Res. Brain Res. Rev.* **36**, 204–212 (2001).
11. Moon, L. D. F. et al. *Nature Neurosci.* **4**, 465–466 (2001).
12. Coumans, J. V. et al. *J. Neurosci.* **21**, 9334–9344 (2001).

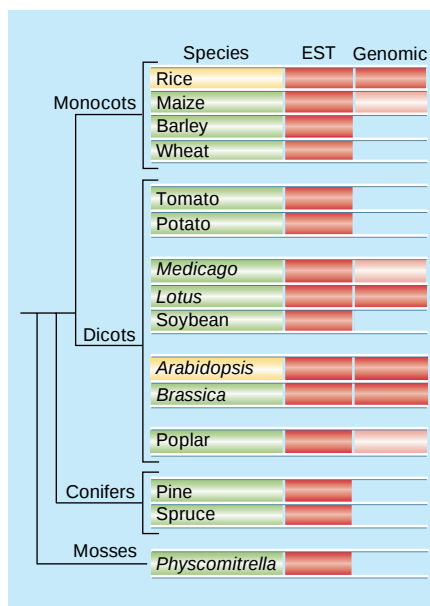


Figure 1 **Progress in plant-genome sequencing.** Large-scale sequencing of representatives of several of the major groups of plants is underway (red bars). Mostly this is through analysis of expressed sequence tags (ESTs), as with pine and moss (*Physcomitrella*). But genome sequencing of *Brassica* species and *Lotus* (a model legume) is ongoing, with various approaches being used, and work on maize, *Medicago* (alfalfa) and poplar will start imminently (pale red bars). The evolutionary relationships shown here are figurative only.

rich regions of the rice genome total about 400 Mbp and contain about 44,700 genes.

The rice genome is much smaller than that of maize (which, at 2,500 Mbp, is about the same size as the human genome), and is tiny compared to the monster barley (twice

human) and wheat (five times human) genomes. From the early estimates of genome size, rice was an obvious target for sequencing, and it probably contains the same gene set in approximately the same gene order⁵ as the other cereal genomes. So it is an ideal genome from which the identity of genes in wheat and maize, the other main cereal crops, can be inferred.

The principal reference genome for the dicots is that of the cress *Arabidopsis thaliana*, the complete sequence of which was published a couple of years ago⁶. Comparison of the predicted rice genes to the *Arabidopsis* genome by one of the groups¹ reveals a limited degree of conserved gene order, extensive chromosomal and whole-genome duplications, and selective gene loss during the evolution of flowering plants. These data provide a fertile field for dating and mapping duplication events in both the monocot and dicot evolutionary lineages, and relating these events to speciation.

The central message from this first comparison of the gene sets of two plants is that they have much in common. Eighty-five per cent of *Arabidopsis* genes have high sequence similarity to rice genes¹, and among these the characteristics of plant genomes established in *Arabidopsis* are further confirmed in the distantly related rice. These data vindicate the strategy of concentrating on systematically defining the functions of *Arabidopsis* genes⁷, using the well developed resources of the large research community that studies this plant.

Draft sequence data always give intriguing hints that require further investigation. In the case of rice, the distribution of repeated sequences is of key interest. The vast size differences in monocot genomes arise primarily by expansion of DNA-repeat families, such as transposons⁸ and other mobile elements that can reproduce copies of themselves throughout the genome. This expansion is thought to occur in interspersed clusters, leaving islands of gene-rich sequence separated by tracts of gene-poor DNA that is generally modified by methylation of certain bases in the large genomes of maize, barley and wheat⁹. Within the grass lineage, the gene order in these islands may be much the same. So the complete, contiguous rice sequence would provide a foundation for gene discovery in all cereals.

Therefore, the next goal is to obtain complete, contiguous sequence that is essentially gap-free and highly accurate. This requires a lot more time and money than draft sequencing, so the task is shared by collaborators in the publicly funded International Rice Genome Sequencing Project (IRGSP). Three of the 12 chromosomes of Nipponbare¹⁰ are now nearly finished, and complete coverage is slated for 2004. This work has been substantially accelerated by the early contribution of sequences from another biotech company, Monsanto¹¹, which com-

prise about 30% of the public sequence. If Syngenta follows this lead, the genome sequence will be completed much sooner, generating a rich resource for gene discovery and key information for understanding the genome dynamics of cereals. But, as reported in various news stories¹², there is controversy over genome publication without full public access to the data.

The catalogue of rice genes presented in these publications^{1,2} complements the *Arabidopsis* data and greatly extends our knowledge of plant genes and genomes. There are around 250,000 species of flowering plant, however, not counting the numerous other plant types such as conifers, ferns and mosses, and having the sequence of two genomes scarcely reflects either the biological richness of the world's plant heritage or its biotechnological potential.

Nevertheless, the overall task of plant genome sequencing is now not as daunting as it might seem. Figure 1 summarizes the present state of play. Expressed sequence tags (ESTs, short sequences of expressed genes) from species as diverse as moss and pine are being gathered. The highly accurate, near-complete rice sequence being generated by the IRGSP, and the *Arabidopsis* sequence that was produced using similar strategies, form the reference genomes for

the monocot and dicot lineages respectively. Consequently, other genomes may not need to be sequenced to these high standards, and whole-genome shotgun sequencing or other rapid and cost-effective strategies can be used on other plants. The genomes of poplar, two legumes, tomato and maize are currently being sequenced using different strategies, and a novel method¹³ for enriching and sequencing gene-rich DNA from the large cereal and grass genomes promises to bring these important targets within reach. All in all, the stage is set for an explosive growth in sequence accumulation from even the most recalcitrant plant species. ■

Michael Bevan is at the John Innes Centre, Norwich NR4 7UH, UK.

e-mail: michael.bevan@bbsrc.ac.uk

1. Goff, S. A. et al. *Science* **296**, 92–100 (2002).
2. Yu, J. et al. *Science* **296**, 79–92 (2002).
3. Chen, M. et al. *Plant Cell* **14**, 537–545 (2002).
4. Wu, J. et al. *Plant Cell* **14**, 525–535 (2002).
5. Gale, M. D. & Devos, K. *Proc. Natl Acad. Sci. USA* **95**, 1971–1974 (1998).
6. The *Arabidopsis* Genome Initiative *Nature* **408**, 796–815 (2000).
7. Chory, J. et al. *Plant Physiol.* **123**, 423–425 (2000).
8. San Miguel, P. et al. *Science* **274**, 765–768 (1996).
9. Moore, G. et al. *Curr. Biol.* **5**, 737–739 (1995).
10. <http://demeter.bio.bnl.gov/Tsukuba02.html>
11. Barry, G. F. *Plant Physiol.* **125**, 1164–1165 (2001).
12. Butler, D. *Nature* **416**, 111–112 (2002).
13. Rabinowicz, P. D. et al. *Nature Genet.* **23**, 305–308 (1999).

Geophysics

The disappearing dipole

Peter Olson

Satellite measurements of the Earth's magnetic field reveal a detailed picture of the circulation in the liquid iron core. The data suggest that the planet could be in the early stages of reversing its magnetic polarity.

The motion of the Earth's liquid core is the geodynamo that generates the planet's magnetic field. Because the electrical conductivity of the molten iron in the core is high, the magnetic flux produced by the dynamo closely follows the fluid motions in the core. Much as oceanographers use dye to trace the ocean circulation, geophysicists can monitor the gradual changes in the magnetic field at the Earth's surface to trace the circulation of the planet's liquid outer core. Magnetic flux lines are transported, rotated, sheared and stretched by flow in the core, inducing slowly varying irregularities in the magnetic field at the Earth's surface (the 'secular variation'). Projecting these magnetic irregularities down to the outer-core boundary reveals the motion of the liquid core at that depth, and provides a real-time image of the geodynamo in action. On page 620 of this issue, Hulot et al.¹ use this technique with satellite measurements of the magnetic field taken 20 years apart to produce high-resolution images of

the outer-core flow. Their results confirm some long-held tenets of dynamo theory — but contradict others.

Important clues about how the geodynamo operates can be found in the detailed structure of the magnetic field at the core–mantle boundary. Consider, for example, the precipitous decline of the Earth's main dipole field over the past 150 years. At its current rate of decrease, the dipole field will vanish early in the next millennium. This is an amazingly rapid change, in view of the fact that the main dipole field has existed for at least the last 3 billion years of the Earth's history. To put the present decline into perspective, suppose the liquid core were suddenly to freeze, so that all circulation stopped immediately. The dipole field intensity would then decrease exponentially through ohmic resistance — but slowly, by only about 1% per century.

High-resolution images of the geomagnetic field taken in 2000 by the Oersted satellite, combined with similar images taken in

1980 by the Magsat satellite, show what is responsible for the rapid dipole decline. Hulot *et al.*¹ have identified patches of reversed magnetic flux concentrated in two regions on the core–mantle boundary. The largest of these patches is beneath the southern tip of Africa. Here, the magnetic field points towards the centre of the Earth, opposite to the normally outward-pointing field of the Southern Hemisphere. Another concentration of reversed-flux patches is found in the north polar region. Growth and poleward migration of the reversed-flux patches can account for almost all of the decrease in the dipole field in the past 150 years².

Numerical models of self-sustaining fluid dynamos³ also indicate that reversed-flux patches are important in magnetic polarity reversals. A self-sustaining fluid dynamo works by positive feedback: the relative motion between the conducting fluid and the ambient magnetic field induces a secondary amplifying field. Positive feedback means that the same dynamo process will also amplify any reversed ambient field, so that localized antidynamos with reversed field can coexist within a larger dynamo. At any instant in time, the geomagnetic field on the core–mantle boundary probably contains several reversed-flux patches, each acting as a localized antidynamo. As the reversed-flux patches merge and grow, the antidynamo might become strong enough to

cause a full-blown dipole polarity reversal.

So are we witnessing the early stage of a polarity reversal in the growth and migration of reversed-flux patches? It is premature to conclude that the present-day downward trend of the dipole field will continue until the field is gone, particularly as the geological record shows that the magnetic field intensity has oscillated in the past without actually reversing its polarity⁴. But the rapidly evolving reversed-flux patches suggest that an attempt at reversal may be underway. In any case, their presence can explain why polarity reversals, once initiated, can happen over only a few thousand years.

The circulation in the outer core also contains some surprises. One of the oldest concepts in geomagnetism is the westward drift. It was originally proposed⁵ by Edmond Halley in 1692 to explain the variation in magnetic directions measured by ships along the trade route from Europe to the Far East. It turns out that Halley's discovery was fortuitous. We now know that the geomagnetic westward drift is anomalously large over that portion of the Earth. In fact, the global-average westward drift is remarkably small — only around 0.1° per year. It varies on a time-scale of decades, because of the oscillation of the core and the mantle; the core rotates slower when the mantle rotates faster, and vice versa, conserving the angular momentum of the Earth⁶.

Significant deviations from uniform westward drift are confined to vortices in the polar regions. Polar vortices are prominent features in other geophysical circulations, most notably in the stratosphere. The polar vortices in the core probably have the same origin as their atmospheric counterparts (that is, lateral temperature gradients) and they may extend to great depths, possibly to the solid inner core. According to Hulot *et al.*¹, the polar vortices are confined within 21° of the poles — the latitude where a cylinder drawn tangentially to the inner-core equator would intersect the core–mantle boundary. This suggests that the inner core may exert topographic control on lateral thermal gradients in the outer core and on the size of the polar vortices.

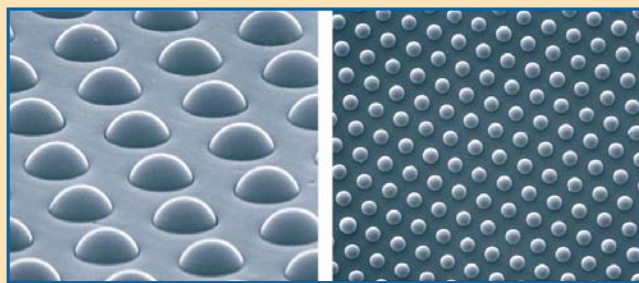
But Hulot *et al.* also find evidence for additional vortex-style flows outside the inner-core tangent cylinder. These vortices may represent the tops of geostrophic convection columns, in which the flow is nearly invariant along the direction of the Earth's rotation axis. A puzzling property of these vortices is their asymmetric distribution: the prograde vortices (with fluid rotation parallel to the Earth's rotation direction) cluster in the Pacific hemisphere, but the retrograde vortices cluster in the Atlantic hemisphere. A longer record of the geomagnetic history would determine whether this asymmetry is a lasting feature of the core field. According

Molecular biology

Alternative arrays

As the sequencing of the human genome nears completion, the full catalogue of encoded genes is coming into focus. In an effort to find out where and when each of these genes may be active in our bodies, researchers often turn to microarrays, which allow the levels of many different DNA or RNA sequences to be measured at once. But existing methods may not be specific or sensitive enough for some tasks, such as studying alternative splicing — a process by which RNA copied from a single gene may be 'cut-and-pasted' in varying ways to produce different messenger RNA (mRNA) molecules and so, ultimately, different proteins. Joanne M. Yeakley and colleagues (*Nature Biotechnol.* **20**, 353–358; 2002) now describe how they used a fibre-optic array to study the alternative splicing of human genes.

A bundle of some 50,000 optical fibres makes up the business end of the array, with one end of the bundle



being etched to allow a thousand or so tiny glass beads, each bearing a specific DNA sequence, to adhere stably (as shown in close-ups here). The positions of individual DNA molecules on the array are first mapped by using fluorescent DNAs of known sequence that attach to matching DNAs on the beads. The pattern of fluorescence is read by sensitive detectors at the other end of the fibre-optic bundle.

To detect mRNAs produced by alternative splicing, Yeakley *et al.* combine the fibre-optic array with a technique dubbed RASL (for 'RNA-

mediated annealing, selection and ligation'). The mRNAs produced in a particular cell type are extracted and mixed with pairs of short DNA molecules, each pair having been chosen to span a presumed splice junction — a sequence at which two RNA segments can be joined by splicing. If the mixture contains an mRNA with a particular splice junction, the corresponding DNA duo binds to that mRNA and the two DNAs can be fused enzymatically. The DNAs are then multiplied, fluorescently labelled, and finally detected by binding to

matching sequences on the beads of the array.

The authors use their approach to look for about 100 individual mRNAs in various human cell types. They find that the system is reliable and sensitive, allowing mRNAs to be detected from fewer than 10 cells. The compactness of the fibre-optic bundles means that an 'array of arrays' — with many bundles arranged in a block — can be constructed, allowing parallel experiments to be carried out. But the method can only pinpoint splice junctions that are predicted from known DNA-sequence information, and understanding complex patterns of alternative splicing will require further analysis. Nonetheless, the approach should prove useful in exploring the ability of alternative splicing to generate functionally diverse proteins in different cells, perhaps in different developmental or physiological conditions.

Richard Turner



100 YEARS AGO

The "red rain" which fell in many parts of Italy and extended as far as Vienna and other central European stations on the evening of March 10, 1901, has been subsequently studied by Prof. N. Passerini, and an account of the phenomena is now given by him in the *Bollettino mensuale* of the Italian Meteorological Society. The phenomenon appears to have travelled slowly from south to north... Prof. Passerini found that the precipitation of the earthy substance was accompanied with very little rain, and a rough analysis showed it to contain about 44 per cent. of fine sand, 32 per cent. of argillaceous matter, 12 per cent. of calcareous matter and about 10 per cent. of organic and volatile substances destroyed by calcination. The red colour was probably due to ferric hydrate... It is suggested that the material deposited in this and other so-called "rains of blood" that have occurred at different times in Italy may probably have been transported by a cyclonic disturbance, and may have had its origin in the equatorial regions of Africa or America. From *Nature* 10 April 1902.

50 YEARS AGO

Recently, while continuing the search for additional Lower Miocene fossil fruits and seeds in the deposits on Rusinga and M'fwangano Islands in Lake Victoria, in order to obtain more data concerning the environmental conditions under which the Miocene ape *Proconsul* lived, we have found some remarkable fossils of insects and other soft-bodied invertebrates. This discovery is of unusual interest, not only because it is believed to be the first of its kind in Africa, but also because of the very unusual state of preservation of the material. Most fossil insect remains — other than those preserved in Oligocene ambers — have been found in a somewhat crushed and distorted condition in laminated rocks. In the present case the fossils, in spite of their soft-bodied nature during life, have retained their natural shape in a quite surprising manner... Some twenty specimens of invertebrates have so far been obtained, and it is anticipated that many more will be found as the work of the British Kenya Miocene Expedition continues. The material at present includes...representatives of Orthoptera, Coleoptera, Blattoidea and, perhaps the most surprising of all, what appears to be a fossil earthworm. L. S. B. Leakey From *Nature* 12 April 1952.

to dynamo theory, helical vortex flows like these should be very efficient in maintaining the geomagnetic field, and indeed helical vortices are prominent in numerical models of the geodynamo⁷.

What are the prospects for capturing even more detailed images of the geodynamo? Improved spatial resolution is unlikely to be the answer, because it is difficult to see smaller-scale behaviour of the core magnetic field through the planet's magnetized crust. But by combining data from satellite missions with measurements of the palaeomagnetic field recorded in lava and sediments⁸, there is potential for extending our picture of the geodynamo over thousands, even millions, of years into the past. ■

Peter Olson is in the Morton K. Blaustein Department of Earth and Planetary Sciences, Johns Hopkins University, Baltimore, Maryland 21218, USA. e-mail: olson@jhu.edu

1. Hulot, G., Eymin, C., Langlais, B., Manda, M. & Olsen, N. *Nature* **416**, 620–623 (2002).
2. Gubbins, D. *Nature* **326**, 167–169 (1987).
3. Glatzmaier, G. A., Coe, R. C., Hongre, L. & Roberts, P. H. *Nature* **401**, 885–890 (1999).
4. Merrill, R. T., McElhinny, M. W. & McFadden, P. L. *The Magnetic Field of the Earth: Paleomagnetism, the Core and the Deep Mantle* (Academic, San Diego, 1996).
5. Halley, E. *Phil. Trans. R. Soc. Lond. A* **16**, 563–578 (1692).
6. Jault, D., Gire, C. & LeMouél, J.-L. *Nature* **333**, 353–356 (1988).
7. Roberts, P. & Glatzmaier, G. A. *Rev. Mod. Phys.* **72**, 1081–1123 (2000).
8. Constable, C. G., Johnson, C. L. & Lund, S. P. *Phil. Trans. R. Soc. Lond. A* **358**, 991–1008 (2000).

Global change

Carbon dioxide goes with the flow

John Grace and Yadvinder Malhi

Measurements of the rate at which carbon dioxide is released from rivers running through tropical forests provide a surprise. They will help in developing an improved picture of the carbon cycle.

Rainforests contain not only trees but also lots of water, largely in the form of river systems. The Amazon is by far the largest such system in the world, contributing 20% of all water flowing from rivers to the ocean. But how this and other great rivers participate in the global carbon cycle is a puzzle — relatively small quantities of carbon are detected in the outflow, yet organic material from the adjacent forest is commonly observed as floating debris. On page 617 of this issue¹, Richey *et al.* provide insight into the watery fate of the organic matter produced by the forest, showing that much of it is released into the atmosphere as CO₂. There is some way to go, however, before we can balance the carbon-budget books.

The study¹ focuses on the central part of the Amazon basin (see the map on page 617). The authors show that the river water contains high concentrations of CO₂, which they infer is being released into the atmosphere at a surprisingly high rate per unit area — a rate that is comparable to that from respiration by soil organisms, and which contributes substantially to the 'decomposition flux' of CO₂ from the basin as a whole.

The recognition that the forest exports carbon, which decomposes in water and is released as CO₂, helps to reconcile some of the conflicting results from studies of the carbon cycle and carbon budget, both in the Amazon and on a global scale. The data concerned come from measurements by eddy covariance^{2,3} (direct measurements of CO₂ flux made above the forest, using towers), analyses of forest inventories⁴ (calculations

from repeated measurements of the number and size of trees in sample plots), and studies of the global atmosphere⁵ (calculations of the geographical distribution of CO₂ sources and sinks made from frequent and precise measurements of concentration in the Earth's atmosphere).

Several studies of eddy covariance suggest that about 5 × 10⁶ g of carbon accumulate per hectare per year in the dry-land — 'terra firme' — forests of the Amazon basin. This is a surprisingly large amount, but to our knowledge only two studies have produced much smaller estimates². Although it is expected that the 'fertilization' that arises from increasing levels of CO₂ in the atmosphere would enhance the uptake of carbon, with photosynthesis exceeding respiration, it seems from both calculations and experiments that this effect is rather small for mature forests⁶. If the high rate of carbon accumulation were to apply to the entire basin, we would have to conclude that the Amazon is a giant carbon sink, perhaps absorbing two-thirds of the world's fossil-fuel emissions. This seems unlikely, and moreover is not consistent with either forest-inventory analyses or global atmospheric studies, which suggest that the Amazonian sink is much smaller. Richey *et al.*¹ show that significant quantities of the carbon assimilated by the forest are almost certainly carried away, and decompose elsewhere. In other words, the eddy-covariance towers situated in the forests do not give a true picture of the overall decomposition.

These, however, are highly complicated issues, and inevitably there are questions to

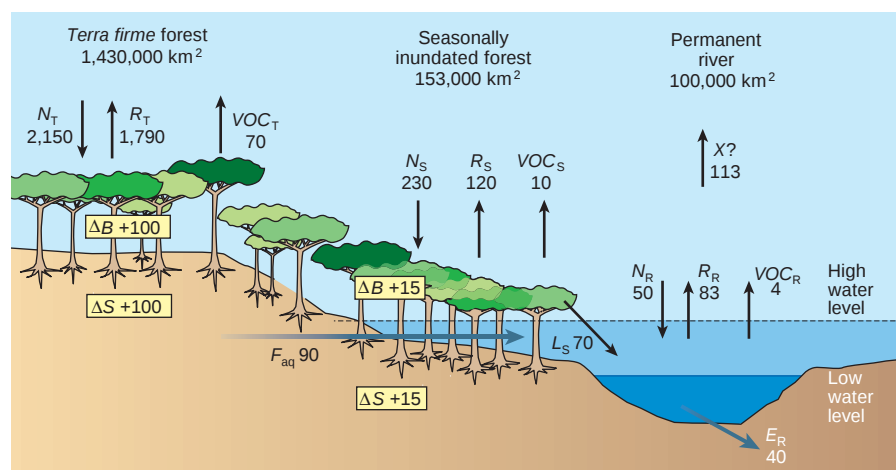


Figure 1 Estimates of carbon flows in the central Amazon basin. As defined by Richey *et al.*¹, this is a region of 1.77 million km², and here it is shown divided into *terra firme* forest (T), seasonally inundated forest (S) and permanent river (R). All units are 10¹² g C yr⁻¹. The area of seasonally inundated forest (153,000 km²) represents the mean area that is flooded over the year. *N*, net primary productivity; *R*, heterotrophic respiration; *VOC*, flux of volatile organic compounds (methane, isoprene, monoterpenes); *F_{aq}*, flow of carbon from forest to river as dissolved CO₂ and dissolved organic carbon; ΔB , net carbon accumulation in biomass; ΔS , net carbon accumulation in soils; *L_S*, direct litter flow from flooded and riverside forests into water; *E_R*, net export of dissolved or inorganic carbon in the Amazon's main stream. *N_T* and *R_T* are derived from ref. 7. *N_S* per unit area is assumed to be the same as *N_P* (the net primary productivity of the *terra firme* forest⁸). ΔB and ΔS are derived from forest inventories and estimates of soil-carbon residence times⁹. *E_R* is scaled down from the estimate for the entire Amazon basin¹. *X* is a residual efflux from the river, which is required to account for the fact that the flux reported by Richey *et al.* is less than the carbon entering the river.

be asked about the new results. For instance, is the authors' method for calculating CO₂ flux reliable? The method is the same as that used by oceanographers, and is relatively simple: flux is calculated as the difference in CO₂ partial pressure between the water and the atmosphere (in this case, measured at 1,800 sites), multiplied by an exchange coefficient. As used by Richey *et al.*¹, the calculation can be expected to yield a conservative estimate of CO₂ flux under most conditions because the exchange coefficient is determined inside floating chambers, without normal air movement, so the true flux may be even higher.

To see the new work in a broader context, the central Amazon basin can be represented as a three-component system: the *terra firme* forest, the seasonally inundated forest and the permanent river. To try to identify the uncertainties in the carbon budget, we have combined Richey and colleagues' data on outgassing of CO₂ by the river with published carbon fluxes over *terra firme* forest⁷ and textbook data on fluxes over inundated forest⁸. The outcome is shown in Fig. 1. We calculate that the residual carbon flux from the river, which must balance the inputs and outputs, is 1.13×10^{14} g C yr⁻¹. This suggests either that the authors' estimate of CO₂ flux is indeed conservative, or that the calculated input of dissolved carbon from *terra firme* forest, estimated by Richey *et al.* as 40% of the total riverine CO₂ flux, is too high.

Even taking the export of carbon into account, there remains a problem in balancing

the books for *terra firme* forests. Net carbon uptake is often measured by eddy covariance as 500 g C m⁻² yr⁻¹, of which about 70 g are accumulated in the biomass and 70 g in the soil carbon pool, 50 g are lost as volatile

compounds, and 60 g are exported to the river. This leaves 250 g C m⁻² yr⁻¹ of carbon uptake as 'missing carbon', which must be accounted for. One possible explanation, which is currently being explored, is that, during the night, CO₂-rich air in the forest canopy drains to valley bottoms and towards rivers.

Finally, to interpret fully the results of Richey *et al.*, we need to know more about the origins of the CO₂ that comes from the river. The sources are likely to be organic matter from *terra firme* forest, from inundated forest and from vegetation on the riverside and in the river itself. Richey *et al.* provide an estimate of the relative importance of these sources, but their figures are highly uncertain. We also need clarification on related questions, such as whether the carbon is young (from decaying litter) or old (from within the soil). Studies using chemical and isotopic tracers are the way forward in this respect.

John Grace and Yadvinder Malhi are at the Institute of Ecology and Resource Management, University of Edinburgh, Edinburgh EH9 3JU, UK.
e-mail: jgrace@ed.ac.uk

1. Richey, J. E., Melack, J. M., Aufdenkampe, A. K., Ballester, V. M. & Hess, L. L. *Nature* **416**, 617–620 (2002).
2. Grace, J. *et al.* *Science* **270**, 778–780 (1995).
3. Malhi, Y. *et al.* *J. Geophys. Res.* **103**, 31593–31612 (1998).
4. Phillips, O. L. *et al.* *Science* **282**, 439–442 (1998).
5. Gurney, K. R. *et al.* *Nature* **415**, 626–630 (2002).
6. Lloyd, J. & Farquhar, G. D. *Funct. Ecol.* **10**, 4–32 (1996).
7. Malhi, Y. *et al.* *Plant Cell Environ.* **22**, 715–740 (1999).
8. Junk, W. *The Central Amazon Floodplain* (Springer, Berlin, 1997).
9. Malhi, Y. & Grace, J. *Trends Ecol. Evol.* **15**, 332–337 (2000).

Immunology

DNA drives autoimmunity

Carola G. Vinuesa and Christopher C. Goodnow

In autoimmune diseases, a person's tissues are destroyed by their own immune system. IgG, a normal component of blood, provokes autoimmune responses when immune cells recognize it as a complex with DNA.

The fundamental task of our immune system is to fight invading microorganisms while sparing normal components of our own tissues. For example, antibodies (also known as immunoglobulins) secreted by immune cells are usually directed towards molecules on pathogens but not those on 'self' tissues; selective binding of antibodies to the pathogens helps to destroy them. Numerous mechanisms, referred to collectively as immunological tolerance, normally safeguard against self-targeted immune responses. When these mechanisms fail, autoantibodies — self-reactive antibodies — are produced. In systemic autoimmune diseases such as rheumatoid arthritis, antibody molecules themselves are frequently the target of autoantibodies, and these anti-immunoglobulin autoantibodies

are known as rheumatoid factors. On page 603 of this issue, Leadbetter and colleagues¹ reveal an unexpected mechanism that explains how immune cells might perceive antibodies as pathogens, provoking the production of autoantibodies.

B lymphocytes — the precursors of antibody-producing cells — must integrate signals from two classes of signal-detecting receptor protein before they proliferate and make antibodies that will be secreted into the bloodstream. This dual requirement is key to the immune system's ability to distinguish between foreign and self 'antigens' (molecules that elicit an immune response)². Signal 1 comes from the binding of antigens directly to specific antigen receptors on B cells, namely cell-surface immunoglobulin molecules with variable antigen-binding

sites. Signal 2 often comes from cytokine proteins and cell-surface growth factors made by a second class of immune cell, helper T cells, that recognize the same antigen. In other words, the production of circulating antibodies in response to many pathogens depends on recognition of those pathogens by both T and B cells. This constraint greatly reduces the risk of auto-antibody production, as T and B cells would have to escape their tolerance to the same self-antigen simultaneously.

For several important classes of micro-organism, however, signal 2 can come from specific components of the pathogens themselves, rather than through T cells. The best understood of these T-cell-independent antigens is lipopolysaccharide, a major constituent of the cell wall of Gram-negative bacteria. The unique lipid group of lipopolysaccharide is a structural motif found in most such bacteria, and is recognized by a surface receptor present on all B cells, Toll-like receptor 4 (TLR4)³. TLR4 signals through an ancient pathogen-sensing pathway that was first discovered in fruit-flies and involves the Myd88 protein and the gene-transcription factor NF- κ B. High concentrations of lipopolysaccharide stimulate the proliferation of all B cells. But at lower concentrations, the simultaneous recognition of lipopolysaccharide by antigen receptors and by TLR4 causes a synergistic signal, which provokes proliferation and production of circulating antibodies selectively by bacteria-specific B cells (Fig. 1, overleaf)⁴⁻⁶.

Leadbetter *et al.*¹ now reveal that, in a mouse model of systemic autoimmune diseases, the simultaneous activation of cell-surface antigen receptors and another Toll-like receptor, TLR9, causes a particular subclass of self-immunoglobulin — IgG2a, part of the IgG class — to be recognized by B cells as if it were a pathogen. This triggers T-cell-independent proliferation of the B cells.

The authors took advantage of a genetically engineered mouse strain in which most B cells bear surface antibodies with low affinity for self-IgG2a. Previous experiments showed that, because this affinity is low, binding of self-IgG2a is not normally enough to trigger signal 1, so the immunoglobulin neither activates the B cells nor makes them tolerant. But when the mice are bred with a strain that is prone to several autoimmune disorders (systemic lupus erythematosus, vasculitis and arthritis), the self-IgG2a in the offspring becomes potentially 'immunogenic', resulting in high concentrations of circulating rheumatoid-factor autoantibodies⁷.

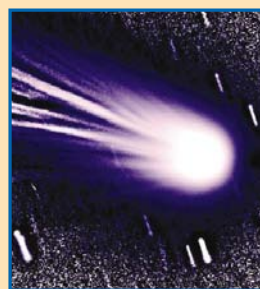
It was not clear why this should happen. But Leadbetter *et al.*¹ find that, in the bloodstream of the autoimmune mice, self-IgG2a accumulates in complexes with self-DNA.

Astronomy

The wanderer returns

Astronomers are celebrating the return of a comet last seen during the seventeenth century. Ikeya-Zhang, named after the Japanese and Chinese amateur astronomers who spotted it on 1 February of this year, is the brightest comet seen from Earth since Hale-Bopp's visit in 1997. Although comets with longer periods may exist, for the time being Ikeya-Zhang has set a new record: its return after 341 years gives it the longest period of any comet so far observed making successive orbits of the Sun.

The historical records of periodicity cannot be interpreted with certainty but they are persuasive. After observing the comet for about 20 days, Brian Marsden of the Smithsonian Astrophysical Observatory in Cambridge, Massachusetts, calculated that the period of the comet is 341 years. This, and the shape of its trajectory, match it to a comet first seen by Polish astronomer Johannes Hevelius in 1661. Hevelius's comet was one of those later used by Edmond Halley — of Halley's comet fame — in his treatise on comet periodicity. Chinese astronomical records also report a comet in the sky in



the years 1320 and 979, which would tally with Ikeya-Zhang's 341-year period, but these sightings may not be reliable.

Comets are essentially dirty snowballs of dust and ice, and are thought to be the remnants of the galactic debris that formed the planets of our Solar System. Measurements of the composition and temperature of the dust and ionized gas trailing behind Ikeya-Zhang are in progress, and will be of particular interest to astronomers. Because of its lengthy periodicity, Ikeya-Zhang makes far fewer trips to the warmer climes closer to the Sun than shorter-period comets such as Halley's (which returns at 76-year intervals), and may have quite a different make-up as a result.

The comet is now around

70 million kilometres from Earth and at its closest approach, on 29 April, it will pass within 60 million kilometres. Since it was spotted a couple of months ago, Ikeya-Zhang has put on quite a display for stargazers. Initially, its tail appeared long and wispy, but now it is short and stubby, so that the comet resembles a shuttlecock. Last week, it seemed to graze the bright 'star' of the giant spiral galaxy, Andromeda. But it wasn't such a close shave — Andromeda is 2.5 million light years away.

Seen from Earth, Ikeya-Zhang is moving over the North Pole and should be visible as a faint smudge to observers in the Northern Hemisphere throughout the night. The best time for viewing is before sunrise, when the comet is about 15 degrees above the horizon (the width of a hand held at arm's length roughly marks out 15 degrees). Although it is visible to the naked eye, its brightness varies, and a pair of binoculars or a small telescope will enhance the show. After 29 April, it will gradually fade from view — and won't return until 2343.

Tom Clarke

The DNA is presumably released during normal or pathological cell death, and (for unknown reasons) fails to be cleared from the bloodstream. TLR9 usually serves as a pathogen sensor by detecting certain motifs (unmethylated CpG base pairs) that are more common in bacterial than mammalian DNA⁸. But these features are also found in certain parts of mammalian genes, and Leadbetter *et al.* show that the ability of self-IgG2a to stimulate an immune response results from its recognition by cell-surface antigen receptors and the simultaneous detection of associated self-DNA by TLR9 on B cells (Fig. 1). When the association is broken, or the TLR9-triggered cellular signalling pathway is disrupted, IgG2a loses this ability.

High levels of rheumatoid factors against IgG are not only found in autoimmune-prone mice. They also occur in people with diseases such as rheumatoid

arthritis and Sjögren's syndrome — in which the immune system attacks the joints and, in Sjögren's syndrome, the tissues that produce tears and saliva — as well as in some patients with lupus erythematosus, which affects many more tissues. Although we do not know how rheumatoid factors are involved in these diseases, a high titre at diagnosis correlates with the severity both of later rheumatoid arthritis and of disease in tissues other than joints, such as blood vessels. Leadbetter *et al.*'s results¹ offer an explanation for how the rheumatoid factors are produced. And they may explain another puzzling observation — that rheumatoid-factor autoantibodies are also generated during normal immune responses to acute bacterial, viral and mycobacterial infections. Foreign DNA, like self-DNA, might form macromolecular complexes with IgG, synergistically triggering IgG-binding antigen receptors and

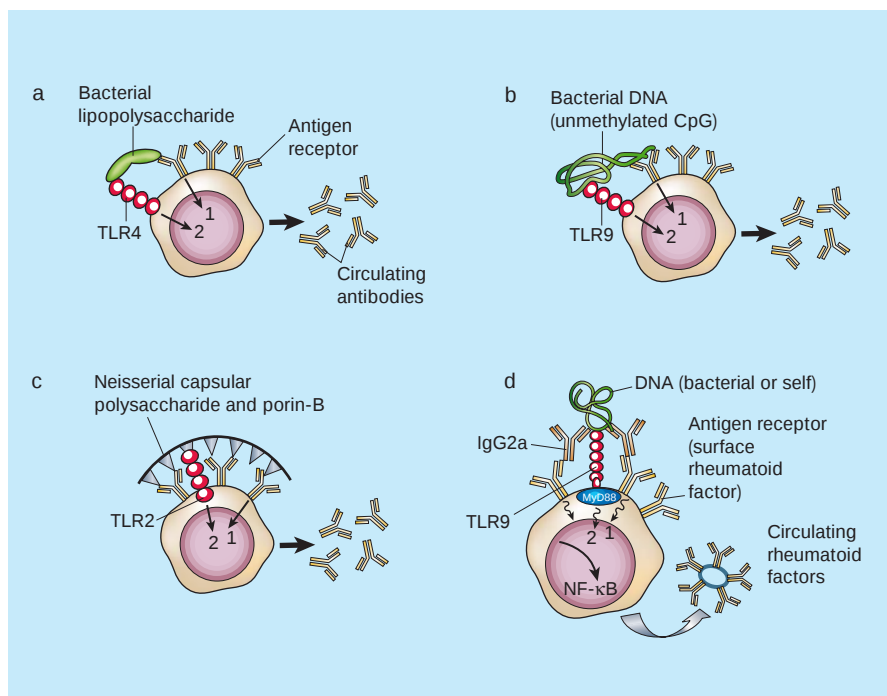


Figure 1 Toll-like receptor proteins in immunity and autoimmunity. a, b, Bacterial lipopolysaccharides (components of the cell wall) or CpG base pairs (a feature more common in microbial than mammalian DNA) can induce T-cell-independent proliferation of, and antibody production by, B cells by stimulating certain Toll-like receptors (TLRs). Simultaneous recognition of the lipopolysaccharide or DNA by specific B cells through cell-surface antigen receptors and TLRs results in synergistic signalling by the receptors. The signals, 1 and 2, lead to the production of bacteria-targeted circulating antibodies. c, Bacterial capsular polysaccharides stimulate specific B cells through their ability to cluster many antigen receptors. For *Neisseria* species, the response can be enhanced¹⁰ by stimulation of TLR2. d, According to Leadbetter *et al.*¹, the binding of complexes containing self-IgG2a and self-DNA (or, presumably, bacterial DNA) to antigen receptors and TLR9 induces B-cell proliferation and the secretion of autoantibodies (rheumatoid factors) against IgG2a.

TLR9 on B cells and prompting the production of autoantibodies against IgG.

Although Toll-like receptors are generally thought to be crucial for sensing pathogens by recognizing structural motifs that distinguish microbes from our own tissues, there are precedents for their binding to self-antigens. As cells die they release various antigens, including heat-shock protein 60 and mitochondrial and nuclear molecules, which can bind to TLR4 and TLR2 on macrophages and trigger these cells to clear up dying cells and thereby help repair damaged tissues⁹. So it is conceivable that B cells, too, may simultaneously recognize these self-antigens by cell-surface antigen receptors and Toll-like receptors, provoking other autoantibodies to nuclear antigens such as self-DNA itself, in diseases that are characterized by prominent tissue damage or the impaired removal of dying cells and DNA. For example, such a process might promote the formation of anti-DNA autoantibodies in systemic lupus erythematosus.

Finally, the finding that Toll-like receptors can drive the production of autoantibodies by their ability to activate B cells single-handedly, without needing T-cell help, should shift some of the clinical focus in

autoimmune diseases towards T-cell-independent processes of antibody production. Leadbetter *et al.*¹ raise the possibility that the drug chloroquine, which inhibits TLR9-mediated signalling, might be effective in some patients with systemic autoimmune diseases. Drugs that specifically target Toll-like-receptor signalling pathways could be promising new treatments for antibody-mediated autoimmune diseases or their complications.

Carola G. Vinuesa and Christopher C. Goodnow are at the Medical Genome Centre, John Curtin School of Medical Research, The Australian National University, Mills Road, PO Box 334, Canberra, ACT 2601, Australia. e-mails: carola.vinuesa@anu.edu.au chris.goodnow@anu.edu.au

1. Leadbetter, E. A. *et al.* *Nature* **416**, 603–607 (2002).
2. Bretscher, P. & Cohn, M. *Science* **169**, 1042–1049 (1970).
3. Poltorak, A. *et al.* *Science* **282**, 2085–2088 (1998).
4. Coutinho, A. & Moller, G. *Scand. J. Immunol.* **3**, 133–146 (1974).
5. DeFranco, A. L., Kung, J. T. & Paul, W. E. *Immunol. Rev.* **64**, 161–182 (1982).
6. Cooke, M. *et al.* *J. Exp. Med.* **179**, 425–434 (1994).
7. Wang, H. & Schlomchik, M. J. *J. Exp. Med.* **190**, 639–649 (1999).
8. Krieg, A. M. *et al.* *Nature* **374**, 546–549 (1995).
9. Li, M. *et al.* *J. Immunol.* **166**, 7128–7135 (2001).
10. Masari, P. *et al.* *J. Immunol.* **168**, 1533–1537 (2002).

Daedalus

Meteoric heating

Many satellites have examined the heat of the atmosphere as a means of predicting the weather. Daedalus now wants to look at the heat of the Earth. The best infrared 'window' through the atmosphere is perhaps 10 micrometres wavelength. Changes in the weather pattern can cause great differences in heat and are not very interesting. Daedalus is looking for small thermal effects. Thermal imaging cameras can be sensitive to a thousandth of a degree centigrade or less. Anomalies of this order on the Earth's surface, sustained local changes from one point to another nearby, demand explanation.

The obvious source of excess local heat is an underground nuclear blast. An unexplained or unacknowledged hotspot in some desert or waste area would show an evasion of the anti-nuclear treaty, and could even give an estimate of its yield. Wider regions of sustained heat might show where coal mines had caught fire. And an earthquake would release energy and warm the ground over a wide area. Indeed, a quake might give a useful thermal signature ahead of time, allowing it to be predicted. Geologists could later map the intensity of the disturbance, independently of seismological records.

The most interesting search, however, would be for meteorite impact sites. On hitting the Earth, a meteorite or comet dumps a lot of thermal energy, which is released for many thousands of years. If it fragments the local crust, heat from within the Earth escapes more copiously thereafter. Many geologists feel that such 'astroblemes' drive evolution. Accordingly, the statistics of meteorite or asteroid descent is a hot topic: will there be another soon? We need to find and date the remains of old ones.

The Barringer crater in Arizona is a famous astrobleme: the site should still be warm. The extinction of the dinosaurs has been blamed on a meteorite that shaped the Yucatán Peninsula of Central America — is that region warmer than it should be? Hudson Bay and James Bay in Canada may be astroblemes — how warm are they? Indeed, all the meteorites that fell into the sea and left unrecorded craters may still be warming the water over the site of their fall. Even human clusters, such as a crowd on a pilgrimage, might be detected.

But the first spot for study must be the Tunguska region of Siberia, where a meteorite fell in 1908. Local heating was intense enough to knock trees over or ignite them. Thermal evidence must still be there.

David Jones

Caterpillar saliva beats plant defences

A new weapon emerges in the evolutionary arms race between plants and herbivores.

Blood-feeding arthropods secrete special salivary proteins that suppress the defensive reaction they induce in their hosts^{1,2}. This is in contrast to herbivores, which are thought to be helpless victims of plant defences elicited by their oral secretions^{3,4}. On the basis of the finding that caterpillar regurgitant can reduce the amount of toxic nicotine released by the tobacco plant *Nicotiana tabacum*⁵, we investigate here whether specific salivary components from the caterpillar *Helicoverpa zea* might be responsible for this suppression. We find that the enzyme glucose oxidase counteracts the production of nicotine induced by the caterpillar feeding on the plant.

Spinnerets are the principal secretory structures of the labial salivary glands of *H. zea*. To determine whether saliva from this caterpillar affects the induced defences of the tobacco plant *in situ*, we prevented salivation by ablating their spinnerets with a heated probe (Fig. 1a, b). This cauterization inhibits salivation, and hence the secretion of salivary enzymes, as we demonstrated by feeding glucose-soaked fibre discs to these caterpillars and then staining the discs to detect hydrogen peroxide (a product of the action of glucose oxidase)⁶. The ablation procedure prevented the release of glucose oxidase without affecting feeding rate.

Caterpillars with ablated spinnerets and caterpillars with intact spinnerets were indi-

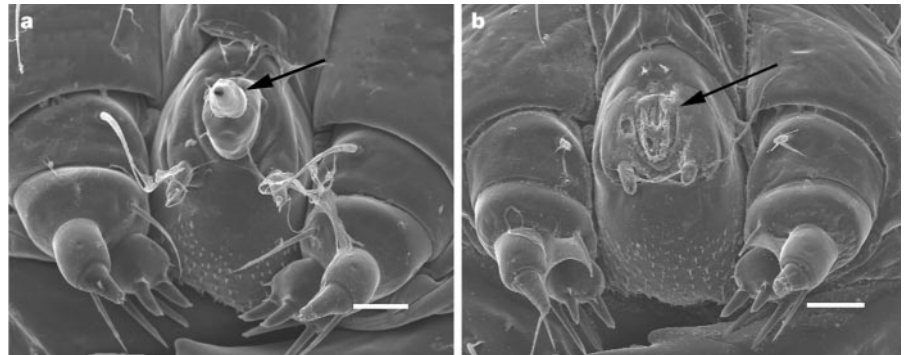


Figure 1 Ablation of the caterpillar (*Helicoverpa zea*) spinneret to prevent production of saliva. **a,b**, Scanning electron micrographs showing the caterpillar labium (arrows) with the spinneret intact (**a**) and ablated (**b**). Scale bars, 100 μm .

vidually caged on the second uppermost fully expanded leaf of an *N. tabacum* plant (one caterpillar per plant) for 24 hours. Three days after feeding, we analysed the damaged leaf for nicotine, an inducible defence compound⁷. Feeding by caterpillars with intact spinnerets reduced foliar nicotine levels by over 26% compared with feeding by caterpillars with ablated spinnerets ($P < 0.05$, Tukey–Kramer test; for details of nicotine quantification, see ref. 7).

We removed circular sections from the leaves of the caterpillar-exposed plants and fed them to *H. zea* neonates. Neonates that were fed on leaves that had been ‘treated’ with caterpillars with intact spinnerets showed significantly increased survival and mean body weights relative to neonates that

were fed on leaves exposed to caterpillars without spinnerets (Mann–Whitney U -test; Fig. 2a, b).

Glucose oxidase is the principal salivary enzyme in *H. zea*, and converts D-glucose and molecular oxygen to D-gluconic acid and hydrogen peroxide^{1,6}. To determine whether this protein could be the salivary component that is responsible for suppressing the plant’s herbivore-induced resistance, we cut one attached leaf per plant with a 1.5-cm-diameter cork borer in order to simulate insect damage. Six holes were uniformly distributed on the second-uppermost expanded leaf.

We then treated individual leaves with one of four preparations: purified, active glucose oxidase (prepared as in ref. 6); unpurified salivary-gland extract; inactivated (autoclaved) purified glucose oxidase; or water. Leaves treated with salivary extract received 20 μg protein in total; each wound received about 10 μl water. Three days later, leaves treated with enzyme or salivary extract contained significantly less nicotine than plants that were treated with water or inactive enzyme (Tukey–Kramer test; Fig. 2c).

We also analysed the survival and growth of second-instar caterpillars on plants treated with either purified glucose oxidase or water. Survival and mean larval weights were significantly greater for caterpillars fed on enzyme-treated leaves than for caterpillars fed on leaves treated with water ($P < 0.05$; Tukey–Kramer test). We also applied the individual reaction products H_2O_2 and gluconic acid to leaf wounds (10 μl 40 mM solution) and found that these products reduced nicotine levels in the leaf relative to a water control by 43.6% and 29.3%, respectively (Tukey–Kramer test; results not shown).

The saliva of herbivorous insects has been overlooked as a factor in overcoming

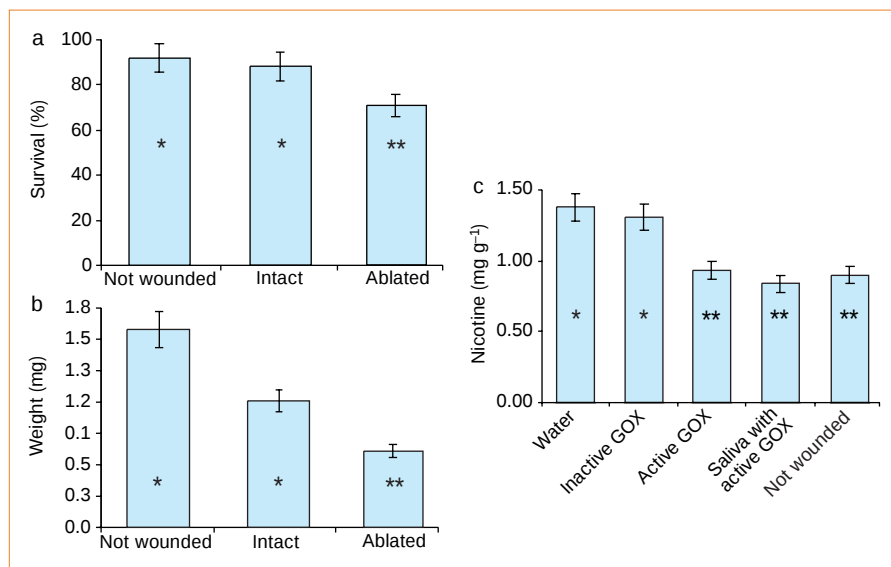


Figure 2 Effect of caterpillar saliva on induced resistance in tobacco plants (*Nicotiana tabacum*). **a**, Proportion of neonates surviving after feeding on leaves previously damaged by sixth-instar caterpillars. **b**, Weights of surviving neonates fed on leaves damaged by sixth-instar caterpillars. **c**, Glucose oxidase (GOX) and salivary-gland extracts suppress nicotine production in leaves that have been wounded to mimic insect damage. Wounded leaves were treated with water, inactive GOX, active GOX, or salivary-gland extracts containing active GOX; the nicotine content of treated leaves was analysed by high-performance liquid chromatography⁷. Asterisks represent significant differences ($P < 0.05$) between treatments; error bars represent mean $\pm$ s.e.

host defences^{1,2}. Our results show that glucose oxidase, one of the principal components of *H. zea* saliva, is responsible for suppressing induced resistance in *N. tabacum*. This enzyme may prevent the induction of nicotine by directly inhibiting the wound-signalling molecule jasmonic acid and/or by antagonizing its interaction with other signalling pathways. As glucose oxidase is produced by a wide variety of caterpillar species^{1,6}, we may have discovered a new feature of the evolutionary arms race between plants and herbivores.

Richard O. Musser*, **Sue M. Hum-Musser†**, **Herb Eichenseer***, **Michelle Peiffer‡**, **Gary Ervin***, **J. Brad Murphy†**, **Gary W. Felton‡**
Departments of *Entomology and †Horticulture,

University of Arkansas, Fayetteville, Arkansas 72701, USA
‡Department of Entomology, Pennsylvania State University, University Park, Pennsylvania 16802, USA
e-mail: gwf10@psu.edu

1. Felton, G. W. & Eichenseer, H. *Induced Plant Defences Against Pathogens and Herbivores* (Am. Phytopathol. Soc., St Paul, Minnesota, 1999).
2. Ribeiro, J. M. C. *Regulatory Mechanisms in Insect Feeding* (Chapman & Hall, New York, 1995).
3. Mattiacci, L., Dicke, M. & Posthumus, M. A. *Proc. Natl Acad. Sci. USA* **92**, 2036–2040 (1995).
4. Alborn, H. T. et al. *Science* **276**, 945–949 (1997).
5. Kahl, J. et al. *Planta* **210**, 336–342 (2000).
6. Eichenseer, H., Mathews, M. C., Bi, J. L., Murphy, J. B. & Felton, G. W. *Arch. Insect Biochem. Physiol.* **42**, 99–109 (1999).
7. Saunders, J. A. & Blume, D. E. *J. Chromatogr.* **205**, 147–154 (1981).

COMMUNICATIONS ARISING

Biodiversity

Suspect evidence of transgenic contamination

Quist and Chapela¹ claim that transgenic DNA constructs have been introgressed into a traditional maize variety in Mexico, and furthermore suggest that these constructs have been reassorted and introduced into different genomic backgrounds. However, we show here that their evidence for such introgression is based on the artefactual results of a flawed assay; in addition, the authors misinterpret a key reference² to explain their results,

concluding that reassortment of integrated transgenic DNA occurs during transformation or recombination.

The discovery of transgenes fragmenting and promiscuously scattering throughout genomes would be unprecedented and is not supported by Quist and Chapela's data¹ — the incorrectly cited work² merely claims that multiple transgenes or transgene fragments can integrate into genetically linked regions of the genome during transformation, and not that they can move around the genome by recombination after integration (W. Pawlowski, personal communication).

The discovery of cauliflower mosaic virus promoter sequences, an element of

transgenic constructs, in the authors' samples is more consistent with F₁ hybridization than introgression. In introgression, a small, polymorphic genomic region is bred into a given variety through repeated backcrossing, so all progeny (or kernels) from an individual in which a (trans)gene has been introgressed will possess that gene. Quist and Chapela report, however, that on the basis of "low amplification" by the polymerase chain reaction (PCR), the transgene is evident only in a small percentage of kernels in each cob, citing a press release that reported a 3–10% abundance of transgenes in similar samples to support their claim.

The authors interpret their inverse PCR (i-PCR) results as evidence of a high frequency of transgene insertion into a range of genomic contexts, inferring from this that introgression events are relatively common and well maintained. However, their i-PCR products all seem to be artefacts of the methodology used. We examined the sequences of the reported i-PCR products (Fig. 1) and found that none contains a reasonable number of the features that would be expected in a legitimate product of amplified genomic DNA flanking the anchor sequence. An i-PCR product derived from the circularization of a single piece of DNA should contain the sequence

Editorial note

In our 29 November issue, we published the paper "Transgenic DNA introgressed into traditional maize landraces in Oaxaca, Mexico" by David Quist and Ignacio Chapela. Subsequently, we received several criticisms of the paper, to which we obtained responses from the authors and consulted referees over the exchanges. In the meantime, the authors agreed to obtain further data, on a timetable agreed with us, that might prove beyond reasonable doubt that transgenes have indeed become integrated into the maize genome. The authors have now obtained some additional data, but there is disagreement between them and a referee as to whether these results significantly bolster their argument.

In light of these discussions and the diverse advice received, *Nature* has concluded that the evidence available is not sufficient to justify the publication of the original paper. As the authors nevertheless wish to stand by the available evidence for their conclusions, we feel it best simply to make these circumstances clear, to publish the criticisms, the authors' response and new data, and to allow our readers to judge the science for themselves.
Editor, *Nature*



Figure 1 Alignment of primers with cauliflower mosaic virus 35S promoter sequences and the ends of inverse polymerase chain reaction (i-PCR) products. The upstream and downstream orientations with respect to the 35S promoter were confused in the original publication¹ and are corrected here. In parentheses: primer name, or GenBank accession number and fragment size. Shaded regions represent the maize genomic sequences with the highest homology (BLAST e-value shown) to the preceding i-PCR fragment. The reverse complements of even-numbered primers are shown for alignment. The footprint of the *EcoRV* restriction enzyme is shown in bold and italicized if in alignment with the site on the 35S promoter. Capitals denote bases that match 35S sequences and are underlined in the regions of the primers; a dashed underline denotes these respective sequences in cases in which they appear at the wrong end and in the wrong orientation on the product. For the hypothetical legitimate i-PCR products, plus signs indicate areas where further transgene DNA would be found.

of the inward and outward primers on the anchor; a footprint of the restriction enzyme in both the anchor and the flanking genomic DNA (reconstituting a restriction site); and the intervening anchor 35S DNA between the primer site and the restriction site. The frequency at which the enzyme used cuts maize DNA indicates that products should average 2,000–4,000 bases.

It would also be expected that some, probably all, of the i-PCR products would contain further transgene sequences adjacent to the 35S promoter. We found that no product contained any transgene sequences. As proof that “introgressed DNA [has] retained its integrity”¹, two sequences are designated by Chapela and Quist as *adh1* sequences, “similar to synthetic constructs ... in transgenic maize ... such as Novartis Bt11”. These sequences show no similarity to the *adh1* intron sequences used in some synthetic constructs, and probably represent retrotransposon DNA.

Several mechanisms may have led to the production of i-PCR artefacts. The sequences of the authors’ primer pairs partially match several genomic sequences, which in turn show high sequence similarity to the amplified fragments (Fig. 1). Spurious products could have been generated depending on the (undisclosed) conditions of the i-PCR; for example, the 3’ ends of the primers used to amplify sequence AF434756 can directly base-pair with the genomic sequence that was amplified (Fig. 1). There was no negative control to address the possibility of i-PCR amplification from maize samples containing no transgenic DNA; furthermore, the restriction enzyme *EcoRV* generates blunt ends on digested DNA, providing no mechanism for preferential ligation of the anchor to digested DNA, rather than to random pieces of sheared DNA.

To consider any of the i-PCR products as legitimate flanking regions requires verification that the anchor is truly adjacent in the genome to the retrieved sequence. Confirmation entails a very simple experiment: PCR using a new outward primer on the anchor and a primer that is specific to the putative adjacent genomic DNA will amplify the same DNA as the i-PCR reaction, as long as the original was legitimate.

An empirical inference from Quist and Chapela’s results is that transgenic corn may be being grown illegally in Mexico, a situation that has already occurred with soybean in Brazil³ and cotton in India⁴. However, the approach used by the authors provides no mechanism for quantifying possible *F*₁ hybridization between traditional and transgenic varieties. The possibility that the 35S signal they detect by PCR in their five samples is due to

contamination cannot be ruled out. No indication is given of the number of repetitions in which each sample produced a positive or negative result, and results from the historical negative control sample are omitted as data not shown, with two lanes of data being excised from the gel in the authors’ Fig. 1.

Transgenic corn may or may not be hybridizing to traditional maize cultivars in Mexico. Whether these events will result in introgression of traits, and whether such introgression could have a negative effect on crop diversity, is pure speculation — so far, there is no evidence of transgenes fragmenting and scattering throughout genomes.

Matthew Metz*, Johannes Fütterer†

*Department of Microbiology, University of Washington, Seattle, Washington 98195, USA
e-mail: metzilla@u.washington.edu

†Institute of Plant Sciences, ETH, CH-8092 Zürich, Switzerland

1. Quist, D. & Chapela, I. H. *Nature* **414**, 541–543 (2001).
2. Pawlowski, W. & Somers, D. A. *Proc. Natl Acad. Sci. USA* **95**, 12106–12110 (1998).
3. Neto, R. B. *Nature* **402**, 344 (1999).
4. Jayaraman, K. S. *Nature* **413**, 555 (2001).

Competing financial interests: declared (see online version).

Published online 4 April 2002; DOI 10.1038/nature738

Maize transgene results in Mexico are artefacts

Quist and Chapela’s conclusion¹ that the transgenes they claim to have detected in native maize in Oaxaca, Mexico, are predominantly reassorted and inserted into a “diversity of genomic contexts” seems to be based on an artefact arising from the inverse polymerase chain reaction (i-PCR) they used to amplify sequences flanking 35S transgenes from cauliflower mosaic virus (CaMV).

After i-PCR, the authors determined eight flanking sequences, two of which (K1 and A3) they claim contained transgenic sequences which they identified as homologous to *adh1*. The *adh1* region, one of the first large regions of the maize genome to be sequenced, is composed mainly of retrotransposons². *Adh1* introns are commonly used to increase expression of transgenes in maize^{3,4}.

Quist and Chapela may have confused a hit to the 160-kilobase (kb) *adh1* genomic region with a hit to an intron of the gene. The *adh1*-homologous regions that flank their amplification products K1 and A3 are located about 40 kb from the *adh1* gene. This sequence is a repetitive element that is also present in two other large maize genomic sequences, and is not an *adh1* intron or coding sequence. In fact, a search of GenBank reveals that K1 is more similar

to an element within the *bronze1* genomic sequence than to the *adh1* sequence.

The eight i-PCR products probably resulted from incorrect PCR priming, because 13 of the 15 base pairs of the authors’ primer iCMV2 can be found in the *bronze1* genomic region (GenBank accession no. AF391808.2) as well as the *adh1* genomic region (GenBank AF123535.1). In addition, the 10 base pairs at the 3’ end of the second primer used, iCMV1, are found within the same *adh1* genomic region. These represent the false priming sites amplified by Quist and Chapela as K1 and A3 (GenBank AF434754, AF434755).

Finally, the final seven base pairs of the primer iCMV2 are identical to a maize Opie retro-element (GenBank U68408.1), which is the third i-PCR sequence amplified, A2 (GenBank AF434756). These false priming sites are found at the boundaries between the primers used and the genomic sequences amplified, suggesting that mobile elements exist in the maize genome that have limited sequence similarity to CaMV. However, this does not show that transgenes are scattered throughout the genome.

Given the following facts — none of the flanking sequences contains an obvious transgene (or any expected CaMV sequence apart from the primers used), i-PCR is prone to generating artefacts⁵, and multiple false priming sites are present in the maize genome — it is likely that the i-PCR sequences are all artefacts and not genuine transgenes.

Southern blots of individual kernels could provide much more reliable information about introgression of transgenes into native populations. Transgenic corn may be being grown illegally in Mexico, but Quist and Chapela’s claim that these transgenes have pervaded the entire native maize genome is unfounded. It is important for information about genetically modified organisms to be reliable and accurate, as important policy decisions are at stake.

Nick Kaplinsky*, David Braun*, Damon Lisch*, Angela Hay††, Sarah Hake*†, Michael Freeling*

*Department of Plant and Microbial Biology, University of California, Berkeley, California 94720, USA

†Plant Gene Expression Center, Agricultural Research Service, US Department of Agriculture, Albany, California 94710, USA

e-mail: nkaplins@nature.berkeley.edu

1. Quist, D. & Chapela, I. H. *Nature* **414**, 541–543 (2001).
2. Sanmiguel, P. et al. *Science* **274**, 765–768 (1996).
3. Callis, J., Fromm, M. & Walbot, V. *Genes Dev.* **1**, 1183–1200 (1987).
4. Chieh, L., Chen, Y., Yu, J. & Shih, D. Y. *J. Food Drug Anal.* **9**, 50–57 (2001).
5. Zimmerman, A., Luthy, L. & Pauli, U. *Lebensm.-Wiss. Technol.* **33**, 210–216 (2000).

Competing financial interests: declared none.

Published online 4 April 2002; DOI 10.1038/nature739

Quist and Chapela reply — Our original publication¹ contained two separate conclusions derived from two methodological approaches. First, using PCR, we detected the presence of three distinct transgenic DNA sequences in maize landraces in Oaxaca, Mexico¹. Second, we attempted to establish the genomic context of transgene insertion using i-PCR. The criticisms raised by Metz and Fütterer and by Kaplinsky *et al.* relate principally to our second statement.

In contrast with the well-established PCR method, i-PCR is an exploratory method that depends on interpretation and the availability of known sequences in databases such as GenBank. We acknowledge that our critics' assertion of the misidentification of sequences labelled with *adh1* intron 1 and with *bronze1* is valid.

The suggestion of mispriming in our i-PCR reaction is also warranted for sequences AF434756 and AF434759 (ref. 1). Significant homology with putative misamplifications is maintained across the length of these fragments, and the CaMV sequence was not recovered. However, this pattern is not found in our other i-PCR sequences. A revealing pattern of discontinuity is found at at least one end of five other sequences, indicating the integration junction between the transgenic DNA and the native host genome. Our critics choose not to recognize this feature in the majority of our i-PCR data. Partial homology with retrotransposon elements in maize is common in primers designed to amplify transposon-like sequences, and is not unique to our primers. Questions concerning the distortion of expected footprints at the DNA-integration junction certainly warrant future work.

The movement of transgenes into new populations and across generations is expected to result in diverse integration patterns^{2–7}. Our findings are compatible with recent studies^{2–6} that characterize transgene/host DNA junctions where rearrangements include interspersions with host or unidentifiable DNA. As altered DNA species should also be an important focus of ecological research, we disagree with our critics who assume that only intact transgenes are worthy of attention in our study.

We agree that PCR-based methods are sensitive and therefore open to artefacts, but strongly disagree that the presence of these artefacts is unavoidable or uncontrollable. The consistent performance of our controls, as reported¹, discounts beyond reasonable doubt the possibility of false positives in our results. Nevertheless, the high sensitivity of the PCR reaction has incited some critics to request a non-PCR-based method to confirm our main statement. To address these challenges, we

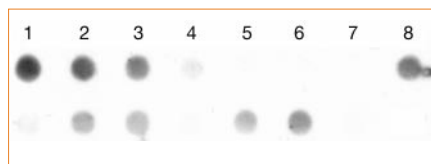


Figure 1 DNA–DNA dot-blot hybridization between maize genomic DNA and a CaMV p-35S probe. Sample numbers coincide with those in ref. 1. Top row: 1, 100% transgenic; 2, 10% transgenic; 3, 5% transgenic; 4, 1% transgenic; 5, 0.5% transgenic; 6, historical maize negative control; 7, water negative control; 8, Diconsa sample K1. Bottom row: 1, criollo sample B1; 2, criollo sample B2; 3, criollo sample B3; 4, criollo sample A1; 5, criollo sample A2; 6, criollo sample A3; 7, Peru maize negative control P1; 8, water negative control.

evaluated the same samples from our original publication¹ using DNA–DNA hybridization. The results of these experiments continue to support our primary statement.

Our analysis of Oaxacan maize is unique for several reasons. First, we wished to document changes that occur within diverse populations of landraces (rather than single varieties or lines), for which no markers, restriction-enzyme digestion maps or linkage analyses have been developed. Second, we could not have predicted which (or how many) specific transgenic constructs (or derivatives) were present in the samples that we analysed. Third, our samples of ground, pooled kernels from individual maize cobs do not represent individual genomes. All of these factors render the application of DNA-hybridization methods difficult. To minimize confusion in interpreting the multiplicity of bands that would have been created by Southern hybridization with our samples, we chose to use dot blotting for our experiments.

We extracted genomic DNA from dry maize kernels¹. Standards containing varying amounts of transgenic material were prepared by mixing flour from our positive control (Bt1) and our historical negative control¹. We blotted and immobilized 10–15 µg of DNA from each sample onto a nylon membrane using a Bio-Dot apparatus (Bio-Rad). We generated a horseradish peroxidase-labelled DNA probe from the same 220-base-pair fragment of the p-35S CaMV promoter that was amplified from our previously reported¹ positive control (Bt1). Hybridization conditions were as follows: 56 °C, 6 ng ml^{–1} DNA probe, 1 hour. Washes were as follows: 3 × 5 min with 0.1 × SSC/0.1% SDS at 56 °C, followed by 3 × 5 min with 2 × SSC at room temperature. Loading homogeneity was confirmed by stripping and rehybridization

of the experimental membrane with the 329-base-pair fragment from the maize-specific *zein* gene¹. Probe labelling, hybridizations and detection were carried out using a North2South kit (Pierce Endogen), according to the manufacturer's specifications.

DNA from four of our six criollo landrace samples, and from the Diconsa sample, hybridized with our CaMV probe (Fig. 1). By using standardized mixtures of transgenic and non-transgenic maize, dot-blot hybridization suggests a ratio of transgenic to non-transgenic kernels in criollo cobs of the order of 1:100, as we had previously suggested¹ and as was confirmed by Mexican government studies¹. This DNA-hybridization study confirms our original detection of transgenic DNA integrated into the genomes of local landraces in Oaxaca.

David Quist, Ignacio H. Chapela

Department of Environmental Science, Policy and Management, University of California, Berkeley, California 94720-3110, USA.

e-mail: ichapela@nature.berkeley.edu

1. Quist, D. & Chapela, I. H. *Nature* **414**, 541–543 (2001).
2. Kohli, A., Gahakwa, D., Vain, P., Laurie, D. A. & Christou, P. *Planta* **208**, 614 (1999).
3. Kumar, S. & Fladung, M. *Mol. Gen. Genet.* **264**, 20–28 (2000).
4. Gorbunova, V. & Levy, A. A. *Nucleic Acids Res.* **25**, 4650–4657 (1997).
5. Windels, P., Taverniers, I., Depicker, A., Van Bockstaele, E. & De Loose, M. *Eur. Food Res. Technol.* **213**, 107–112 (2001).
6. Pawlowski, W. P. & Somers, D. A. *Proc. Natl Acad. Sci. USA* **95**, 12106–12110 (1998).
7. Register, J. C. *et al. Plant Mol. Biol.* **25**, 951–961 (1994).

Competing financial interests: declared none.

Published online 4 April 2002; DOI 10.1038/nature740

correction

Reward value of attractiveness and gaze

K. K. W. Kampe, C. D. Frith, R. J. Dolan, U. Frith
Nature **413**, 589 (2001)

Reward-related responses have been registered in animal brains mostly in the ventral half of the striatum, from the nucleus accumbens to the pallidum. Considering the location of the response to attractive faces we describe, the observed activation was large and its spatial extent was not clear from Fig. 2, although we inferred that the ventral striatum was involved. From the plane shown, this activation more accurately extended ventrally into the striatum, specifically into the pallidum; the nucleus accumbens proper was not activated. Dorsally, the activation extended into the anterior thalamus (as shown in Fig. 2). Our conclusions that the attractiveness of faces is processed in brain regions involved in evaluating the reward value of stimuli, and that this processing depends on gaze direction, are unaltered.

brieF communications is intended to provide a forum for both brief, topical reports of general scientific interest and technical discussion of recently published material of particular interest to non-specialist readers. Priority will be given to contributions that have fewer than 500 words, 10 references and only one figure. Detailed guidelines are available on *Nature's* website (www.nature.com) or on request from nature@nature.com

Chromatin–IgG complexes activate B cells by dual engagement of IgM and Toll-like receptors

Elizabeth A. Leadbetter*, Ian R. Rifkin†, Andreas M. Hohlbaum*, Britte C. Beaudette*, Mark J. Shlomchik‡ & Ann Marshak-Rothstein*

* Department of Microbiology, and † Department of Medicine, Boston University School of Medicine, Boston, Massachusetts 02118, USA

‡ Departments of Laboratory Medicine and Immunobiology, Yale University School of Medicine, New Haven, Connecticut 06520, USA

Autoreactive B cells are present in the lymphoid tissues of healthy individuals, but typically remain quiescent. When this homeostasis is perturbed, the formation of self-reactive antibodies can have serious pathological consequences. B cells expressing an antigen receptor specific for self-immunoglobulin- γ (IgG) make a class of autoantibodies known as rheumatoid factor (RF). Here we show that effective activation of RF⁺ B cells is mediated by IgG2a–chromatin immune complexes and requires the synergistic engagement of the antigen receptor and a member of the MyD88-dependent Toll-like receptor (TLR) family. Inhibitor studies implicate TLR9. These data establish a critical link between the innate and adaptive immune systems in the development of systemic autoimmune disease and explain the preponderance of autoantibodies reactive with nucleic acid–protein particles. The unique features of this dual-engagement pathway should facilitate the development of therapies that specifically target autoreactive B cells.

Patients afflicted with systemic lupus erythematosus (SLE) and other systemic autoimmune diseases produce a wide range of autoantibody specificities. Very frequently these autoantibodies bind to chromatin or other subcellular nucleic acid–protein particles¹. MRL-lpr mice, a well-studied murine model of SLE and rheumatoid arthritis, also produce exceedingly high titres of IgG anti-IgG RF^{2,3}. This RF response has provided a highly relevant transgenic model for the study of autoantibody regulation. B cells from AM14 transgenic mice express an antigen receptor specific for IgG2a^{4/5}, originally captured as a hybridoma product from the spleen of a diseased MRL-lpr mouse⁴. The AM14 allotype specificity and relatively low affinity for monomeric IgG2a are typical of the disease-associated RF repertoire⁵. In wild-type mice, AM14 RF⁺ B cells develop normally and remain functionally naive⁶. However, on an autoimmune-prone lpr background of the cognate allotype, they become activated, proliferate and secrete autoantibody⁷. Several factors are likely to contribute to the activation of AM14 B cells in lpr mice, not the least of which is the status of the autoantigen IgG2a.

Activation of RF⁺ B cells by immune complexes

We have previously shown that AM14 RF⁺ B cells can be activated *in vitro* by IgG2a^{4/5} isolated from the sera of autoimmune mice, but not by comparable levels of IgG2a present in sera obtained from wild-type mice⁸. RF⁺ B cells also proliferate strongly in response to affinity-purified IgG2a¹ monoclonal antibodies specific for nucleosomes, a self-antigen, whereas IgG2a^{4/5} monoclonal antibodies specific for haptens or other foreign antigens produced little, if any, response⁸. These studies suggested that immune complexes formed between IgG2a nucleosome-specific monoclonal antibodies and chromatin fragments released from co-cultured cells⁹ could effectively activate RF⁺ B cells, whereas monomeric anti-hapten IgG2a antibodies could not. To further test this premise, DNase was added to the assay medium as a means of disrupting the putative immune complex. The vigorous proliferative response normally elicited by the purified anti-nucleosome monoclonal antibody PL2-3 (ref. 10) or by autoimmune serum was dramatically reduced (Fig. 1). Nonspecific toxicity of the DNase was not a factor because the responses to F(ab')₂ anti-IgM and lipopolysaccharide (LPS) were

unaffected (Fig. 1). These data confirm that RF⁺ B cells respond to IgG2a immune complexes that contain DNA, not to monomeric IgG2a antibodies alone.

If activation of RF⁺ B cells by IgG2a immune complexes resulted simply from more-effective cross-linking of the antigen receptor than would be possible with monomeric IgG2a, then conventional complexes of anti-hapten antibodies and hapten protein should also strongly stimulate RF⁺ B cells. To address this issue, immune complexes were prepared by pre-incubating IgG2a monoclonal anti-trinitrophenol (TNP) antibodies with varying concentrations of TNP-BSA (bovine serum albumin); complex formation was confirmed by demonstrating an increase in C1q binding activity (Fig. 2a). As expected, immune complexes of C4010, an IgG2a^b anti-TNP monoclonal antibody, failed to stimulate RF⁺ B cells at all anti-TNP/TNP-BSA ratios. Surprisingly, immune complexes of Hy1.2, an IgG2a^a anti-TNP monoclonal antibody, elicited only a very modest proliferative response relative to that elicited by anti-

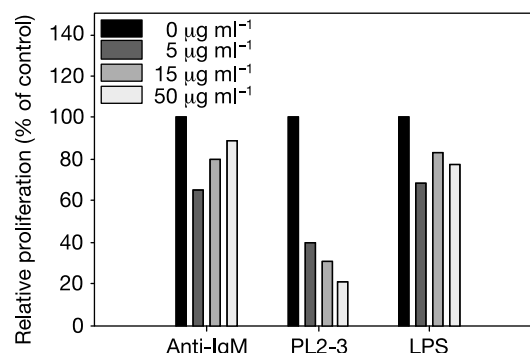


Figure 1 Anti-nucleosome antibody stimulation of RF⁺ B cells is DNase sensitive. AM14 RF⁺ spleen cells were pre-incubated with various concentrations of DNase I (0, 5, 15 or 50 $\mu\text{g ml}^{-1}$) for 15 min before the addition of goat anti-mouse IgM F(ab')₂, PL2-3 (IgG2a¹ anti-nucleosome antibody), or LPS. Results are expressed as the percentage of the maximum response to each ligand in the absence of DNase.

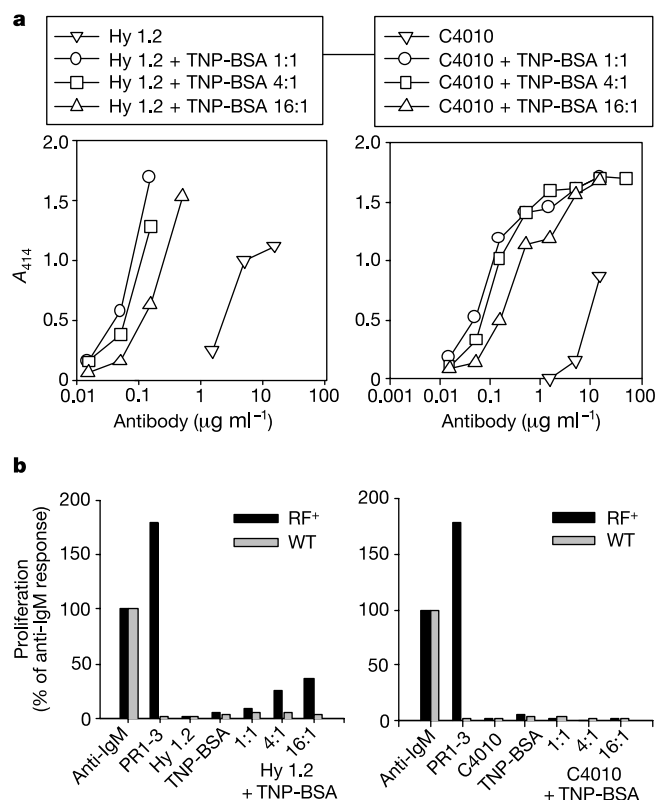


Figure 2 Anti-TNP-TNP-BSA immune complexes fail to efficiently stimulate proliferation of RF⁺ B cells. The anti-TNP monoclonal antibodies Hy1.2 (IgG2a^b) and C4010 (IgG2a^b) were mixed with varying concentrations of TNP-BSA (50 $\mu\text{g ml}^{-1}$, circle; 12.5 $\mu\text{g ml}^{-1}$, square; 3.1 $\mu\text{g ml}^{-1}$, triangle) to form immune complexes with antibody:protein ratios of 1:1, 4:1 and 16:1, respectively. **a**, Immune complex formation was confirmed by a shift in C1q binding activity relative to uncomplexed antibodies (inverted triangle). A_{414} , absorbance at wavelength 414 nm. **b**, The ability of the anti-TNP-TNP-BSA immune complexes described above to stimulate proliferation of RF⁺ B cells was compared with the stimulation induced by the anti-nucleosome antibody PR1-3, uncomplexed Hy1.2, C4010, or 50 $\mu\text{g ml}^{-1}$ TNP-BSA.

nucleosome antibody-containing immune complexes. This discrepancy suggested that the extent of proliferation of RF⁺ B cells evoked by an immune complex depends on both the antibody and the nature of the (auto)antigen.

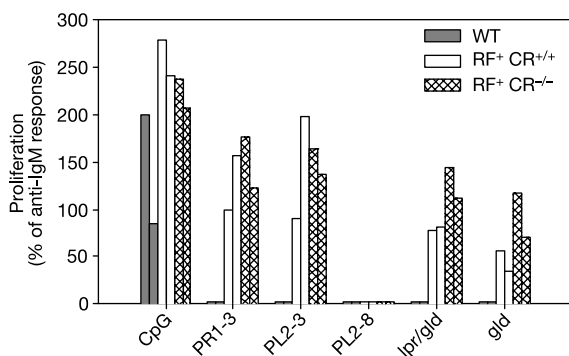


Figure 3 Stimulation of RF⁺ B cells by autoantibody-autoantigen immune complex is not dependent on the complement receptor. Spleen cells from each of two wild-type (WT) control, RF⁺ CR^{+/+} and RF⁺ CR^{-/-} mice were stimulated with goat anti-mouse IgM F(ab')₂, CpG S-ODN 1826, the anti-nucleosome antibodies PR1-3 (IgG2a^b), PL2-3 (IgG2a^b) or PL2-8 (IgG2b), or 3% serum from representative autoimmune mice (Ipr/gld and gld).

CR1/2-independence of RF⁺ B-cell activation

A possible explanation for the marked difference in stimulatory capacity of the anti-nucleosome and anti-TNP immune complexes was that only the former could synergistically engage both the B-cell receptor (BCR) and a second receptor on the B-cell surface. One potential candidate receptor was the complement receptor CD21, because complement components have been shown to bind auto-antigen immune complexes and could, in theory, serve to effectively co-engage CD19/CD21 and the BCR¹¹. However, when the AM14 heavy and light chain transgenes were bred onto mice deficient for Cr2 (ref. 12), RF⁺ B cells from the CR1/2-deficient and CR1/2-sufficient littermates responded comparably to IgG2a anti-nucleosome monoclonal antibodies and autoimmune sera (Fig. 3). B cells from control RF⁻ littermates failed to respond to any form of IgG2a, confirming the specific nature of the ligands used in this study. A mitogenic hypomethylated CpG oligodeoxynucleotide (ODN) was included as a control stimulus. These data demonstrate that complement receptors are not required for the RF⁺ B-cell response to autoantibody-autoantigen immune complexes.

Critical role for a MyD88-dependent receptor

Other potential candidate receptors included members of the TLR family. These pattern-recognition receptors were first described in *Drosophila*, where they were shown to trigger the release of anti-fungal peptides¹³. Subsequently identified mammalian homologues were found to recognize a series of conserved microbial products (such as, LPS, microbial lipoproteins, and hypomethylated CpG DNA) referred to as pathogen-associated molecular patterns (PAMPs)^{14,15}. Initial studies focused on the effects of microbial products on cells of the innate immune system, where they were found to stimulate the release of a wide range of inflammatory mediators¹⁵. However, it was later shown that, in addition to exogenous microbial ligands, TLRs could also recognize endogenous ligands released from damaged or stressed mammalian cells¹⁶. All known mammalian TLRs signal through the adapter protein MyD88 (ref. 15), although in the case of TLR4, an additional MyD88-independent pathway has been described¹⁷. To investigate the potential role of TLRs in responses of RF⁺ B cells to autoantibody-autoantigen immune complexes, we crossed the AM14 transgenes onto a MyD88-deficient (MyD88^{-/-}) background¹⁸. B cells from RF⁻ MyD88^{+/+} (wild type), RF⁺ MyD88^{+/+}, and RF⁺ MyD88^{-/-} offspring, identified by a combination of polymerase chain reaction (PCR) and fluorescence-activated cell sorting (FACS) analysis (Fig. 4a, b), were then assessed for their ability to respond to anti-IgM, CpG S-ODN, and LPS, as well as to a panel of anti-nucleosome monoclonal antibodies and autoimmune sera. As expected, B cells from the MyD88-deficient mice responded normally to anti-IgM, but were unable to respond to stimulatory CpG DNA¹⁹. Their response to LPS was much lower than that of wild-type mice, but still detectable, consistent with partial activation through a MyD88-independent mechanism^{17,39} (Fig. 4c). Most dramatically, B cells from the RF⁺ MyD88-deficient mice were completely unresponsive to the anti-nucleosome monoclonal antibodies PR1-3 or PL2-3 or to any of the autoimmune sera that effectively stimulated RF⁺ MyD88^{+/+} B cells (Fig. 4c). These results clearly demonstrate that autoantibody-autoantigen immune complexes are particularly stimulatory for RF⁺ B cells because they synergistically engage both the BCR and a MyD88-dependent receptor.

Hypomethylated CpG motifs are a common feature of bacterial DNA, but they are also present in mammalian DNA promoter elements²⁰. Because the B-cell response to CpG S-ODN is mediated through TLR9 (ref. 21), it seemed plausible that chromatin-containing immune complexes might stimulate RF⁺ B cells by co-engaging TLR9. The TLR9-mediated response to CpG S-ODN is distinguished from other known TLR signalling pathways by its

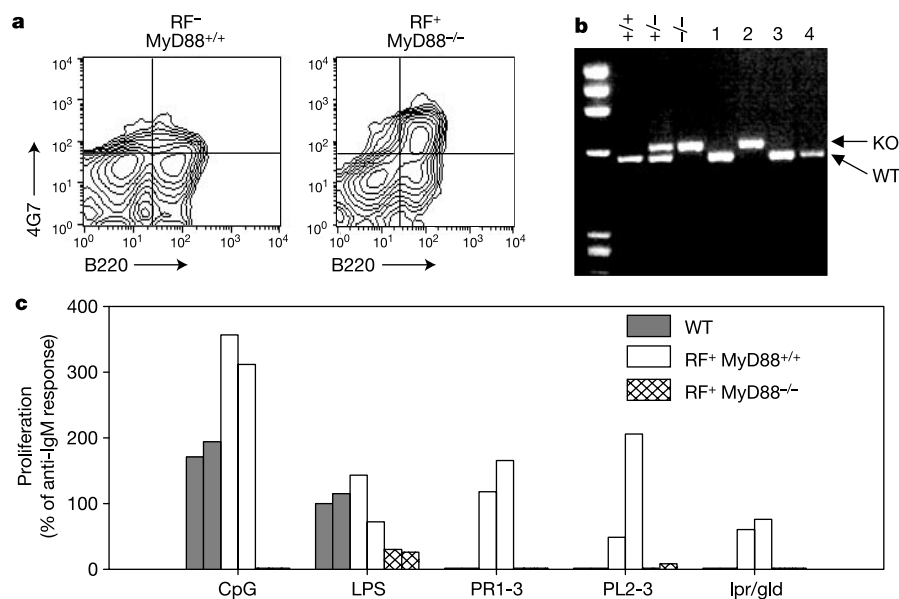


Figure 4 Stimulation of RF⁺ B cells by autoantibody–autoantigen immune complex is dependent on MyD88. **a**, T-depleted spleen cells from wild-type and RF⁺ MyD88^{-/-} mice were stained with B220 and an idiotype specific antibody, 4G7. **b**, MyD88 wild-type or knockout mice were identified by PCR. **c**, Spleen cells from each of two wild-type (WT), RF⁺ MyD88^{+/+} and RF⁺ MyD88^{-/-} mice were stimulated with anti-IgM F(ab')₂, CpG S-

ODN 1826, LPS, anti-nucleosome antibodies PR1-3 or PL2-3, or 3% lpr/gld serum. Total c.p.m. for the anti-IgM-stimulated cultures were 83,414/39,049 (WT), 93,126/61,315 (RF⁺ MyD88^{+/+}), and 39,826/57,484 (RF⁺ MyD88^{-/-}). Results are expressed as the percentage of the anti-IgM response and are representative of four separate experiments.

presumed requirement for endosome acidification and/or maturation as determined by sensitivity to chloroquine and ammonium chloride^{22,23}. Concanamycin B and bafilomycin A are specific inhibitors of the V-type ATPase responsible for acidification of endosomes²⁴. To evaluate the role of TLR9 in the activation of RF⁺ B cells, we determined the effect of chloroquine, concanamycin B, bafilomycin A and ammonium chloride on the stimulatory capacity of monoclonal antibody PL2-3 and known TLR2 (lipopeptide, porin B)²⁵, TLR4 (LPS) and TLR9 (CpG S-ODN 1826) ligands. As expected, these inhibitors had little effect on the response to the TLR2 and TLR4 ligands. Notably, all four agents markedly blocked the response of RF⁺ B cells to monoclonal antibody PL2-3 (Fig. 5a, b, and data not shown). The link to chloroquine is intriguing, because chloroquine is an effective treatment for autoimmune diseases such as rheumatoid arthritis and SLE^{26,27}. Our data raise the possibility that its therapeutic effects may be due in part to its ability to interfere with TLR-mediated signals that contribute to autoantibody production or to the production of pro-inflammatory mediators.

The B-cell response to the stimulatory CpG S-ODN 1826 can also be blocked by a group of closely related inhibitory CpG S-ODN such as S-ODN 2088 (ref. 28). We found that S-ODN 2088 profoundly inhibited the proliferative response to S-ODN 1826 but had no effect on the response to anti-IgM or our TLR2 and TLR4 ligands (Fig. 5c). Most significantly, S-ODN 2088 also markedly blocked the response of RF⁺ B cells to monoclonal antibody PL2-3. Overall, these data strongly implicate endosomal processing and/or engagement of an endosome-associated TLR, most probably TLR9, in activation of RF⁺ B cells.

Discussion

It has been clearly shown that autoreactive B-cell clones can undergo isotype switching and somatic mutation, similar to the T-cell-dependent features of conventional B-cell responses to foreign antigens²⁹. However, other studies have suggested that autoreactive B cells may segregate from conventional B cells in the peripheral lymphoid tissues: autoreactive B cells tend to localize in the marginal zone of the splenic white pulp³⁰, whereas conventional

B cells 'home' to the follicular regions. In addition, in some instances, expansion and somatic mutation of autoreactive B cells can take place outside the germinal centre³¹. These discrete localization patterns may reflect the preferential accumulation of stimulatory autoantibody–autoantigen immune complexes at these sites. Alternatively, the relatively unique homing pattern may result from distinct chemokine receptor profiles elicited in response to the combined effects of BCR–TLR engagement. Future studies will need to compare the functional properties of B cells activated through BCR cross-linking alone from B cells activated through BCR–TLR dual engagement.

Autoantibodies that recognize DNA or nucleosomes are the defining and most prevalent specificity in SLE patients and in murine lupus models of SLE, whereas RFs are characteristic of a subset of SLE patients, autoimmune-prone lpr mice, and rheumatoid arthritis patients. Remarkably, these same autoantigens are the targets in several murine autoimmunity models generated by mutations that disrupt lymphocyte homeostasis or autoantigen metabolism^{32–35}. The reason for the predominance of these particular specificities has long been sought. Our data provide an explanation, namely the potent synergistic interaction of BCR–TLR signalling events mediated by chromatin-containing immune complexes. In our experimental system, dependent on endosome- or lysosome-localized TLR9 (ref. 23), it is reasonable to postulate that engagement of the BCR by an autoantibody–autoantigen immune complex triggers the endocytosis of immune complex-associated antigen that then results in the highly efficient delivery of chromatin fragments to endosome-associated TLR9. In contrast to published studies^{36,37}, we have been unable to activate B cells with either culture supernatants or chromatin fragments alone. Whether other strains of mice will prove more responsive to our fragments remains to be determined.

Beyond the current experimental model, the principle of BCR–TLR dual engagement has wide-ranging implications for autoimmunity in general. Model ligands such as haptenated LPS and LPS-coupled sheep red blood cell can also co-signal BCRs and TLRs and are remarkably potent and specific for the B cells that have the relevant receptors³⁸. Activation in this mode is therefore likely to be

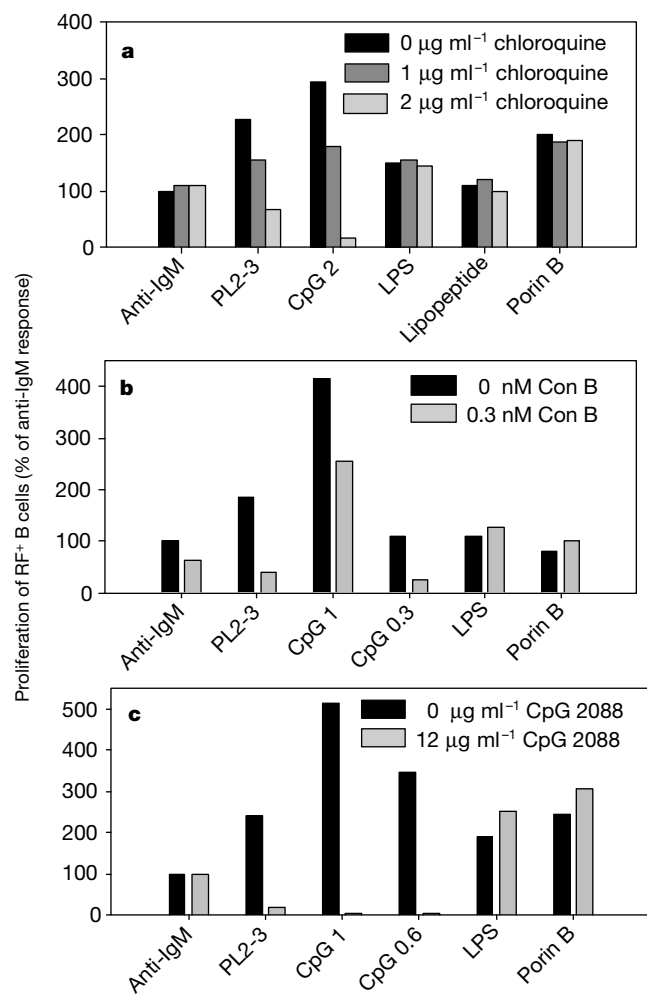


Figure 5 Stimulation of RF⁺ B cells by autoantibody–autoantigen immune complexes can be blocked by inhibitors of the TLR9 signalling pathway. RF⁺ B cells were pre-incubated for 15 min with chloroquine (**a**), pre-incubated for 2 h with concanamycin B (**b**), or pre-incubated for 30 min with inhibitory CpG S-ODN 2088 (**c**), before the addition of the stimulatory ligands anti-IgM F(ab')₂, 5–50 $\mu\text{g ml}^{-1}$ PL2-3, 0.3–2.0 $\mu\text{g ml}^{-1}$ CpG S-ODN 1826, LPS, lipopeptide, or porin B. Results are expressed as a percentage of the anti-IgM response in the absence of inhibitor. The data in **a** and **c** are representative of two to four experiments, whereas **b** is the mean of two experiments.

a fundamental event in the loss of peripheral B-cell tolerance in a wide variety of settings; other autoantigens may signal through TLRs other than TLR9. Overall, the data establish a critical role for endogenous TLR ligands in the aberrant activation of the adaptive immune system in autoimmunity and explain why the autoantibody repertoire is often skewed towards the recognition of sub-cellular nucleic acid–protein particles¹. The ability of microbial PAMPs to engage TLR and/or upregulate TLR expression and thus create a synergy with autoantibody–autoantigen immune complexes might explain the association between infection and disease flares. Most importantly, the identification of signalling components and/or inhibitors specific for this pathway may eventually lead to the development of therapies that specifically target auto-reactive B cells.

Methods

Mice

MRL^{+/+} AM14 BCR Tg mice, described previously^{6,8}, were crossed to Cr2-deficient mice provided by M. Carroll, and the F₁ offspring were intercrossed to generate AM14 Cr2-deficient and Cr2-sufficient control mice. MyD88^{−/−} mice, originally produced by S. Akira¹⁸ and provided through D. Golenbock, were crossed to AM14 MRL^{+/+} mice to

generate AM14 MyD88^{−/−} and control littermate offspring. The RF⁺ offspring were initially identified by PCR, and their identity was confirmed by flow cytometric analysis of peripheral blood lymphocytes or spleen cells, using the 4-G7 monoclonal anti-idiotypic as described⁸. Complement receptor genotype was determined by flow cytometry using the FITC-conjugated 7G6 antibody (Pharmingen). MyD88 genotype was determined by PCR using the primers: MyD88F (5'-TGG CAT GCC TCC ATC ATA GTT AAC C-3'), MyD88R (5'-GTC AGA AAC AAC CAC CAC CAT GC-3') and neoR (5'-ATC GCC TTC TAT CGC CTT CTT GAC G-3') (MWG Biotech) to yield wild-type and knockout products of about 550 and 750 bp, respectively.

Cell culture and reagents

Spleen cell preparations were depleted of T cells and cultured with the appropriate ligands for 40–48 h. In some experiments, B cells were pre-incubated with CD40L as described⁸. Proliferation was assessed by [³H]thymidine incorporation during the final 16 h of culture. Data are presented as the mean percentage of the anti-IgM response from triplicate cultures. The percentage of the anti-IgM response was calculated according to the formula: (c.p.m. experimental condition – c.p.m. CD40L alone)/(c.p.m. anti-IgM – c.p.m. CD40L alone) × 100.

Ligands included goat anti-mouse IgM F(ab')₂ (15 $\mu\text{g ml}^{-1}$, Jackson Immuno Research); the nucleosome-specific monoclonal antibodies PR1-3 (IgG2a¹), PL2-3 (IgG2a¹), and PL2-8 (IgG2b) (all at 50 $\mu\text{g ml}^{-1}$), provided by M. Monestier¹⁰; 10 $\mu\text{g ml}^{-1}$ LPS (Sigma); 0.3–2 $\mu\text{g ml}^{-1}$ stimulatory CpG S-ODN 1826 (ref. 22) (Oligo's Etc.), 10 $\mu\text{g ml}^{-1}$ *Neisseria meningitidis* porin B (provided by L. Wetzler); 10 $\mu\text{g ml}^{-1}$ synthetic lipopeptide Pam₃Cys-Sk₄ (G. Jung, provided by D. Golenbock); and the anti-TNP monoclonal antibodies Hy1.2 (IgG2a¹) and C1040 (IgG2a^b) (both at 50 $\mu\text{g ml}^{-1}$) complexed to varying concentrations of TNP-BSA. In some experiments, DNase I (type IV) (Sigma), chloroquine (Sigma), concanamycin B (provided by H. Ploegh), bafilomycin A (Sigma), or CpG S-ODN 2088 (ref. 28) (Oligo's Etc.) were added to the cultures 15 min to 2 h before the addition of the ligands. All antibody and ODN preparations were shown to be free of endotoxins by Limulus Amebocyte Lysate ELISA (Bio-Whittaker).

Received 20 November 2001; accepted 18 February 2002.

1. Tan, E. Antinuclear antibodies: diagnostic markers for autoimmune diseases and probes for cell biology. *Adv. Immunol.* **44**, 93–151 (1989).
2. Theofilopoulos, A. N. et al. Association of *lpr* gene with graft-vs-host disease-like syndrome. *J. Exp. Med.* **162**, 1–18 (1985).
3. Wolfowicz, C. B., Sakorafas, P., Rothstein, T. L. & Marshak-Rothstein, A. Oligoclonality of rheumatoid factors arising spontaneously in *lpr/lpr* mice. *Clin. Immunol. Immunopathol.* **46**, 382–395 (1988).
4. Shlomchik, M. J., Zharhary, D., Camper, S., Saunders, T. & Weigert, M. A rheumatoid factor transgenic mouse model for autoantibody regulation. *Int. Immunol.* **5**, 1329–1341 (1993).
5. Jacobson, B. A. et al. An isotype switched and somatically mutated rheumatoid factor clone isolated from a MRL-*lpr/lpr* mouse exhibits limited intraclonal affinity maturation. *J. Immunol.* **152**, 4489–4499 (1994).
6. Hannum, L. G., Ni, D., Haberman, A. M., Weigert, M. G. & Shlomchik, M. J. A disease-related rheumatoid factor autoantibody is not tolerized in a normal mouse: implications for the origins of autoantibodies in autoimmune disease. *J. Exp. Med.* **184**, 1269–1278 (1996).
7. Wang, H. & Shlomchik, M. J. Autoantigen-specific B cell activation in Fas-deficient rheumatoid factor immunoglobulin transgenic mice. *J. Exp. Med.* **190**, 639–649 (1999).
8. Rifkin, I. R. et al. Immune complexes present in the sera of autoimmune mice activate rheumatoid factor B cells. *J. Immunol.* **165**, 1626–1633 (2000).
9. Emlen, W., Holers, V. M., Arend, W. P. & Kotzin, B. Regulation of nuclear antigen expression on the cell surface of human monocytes. *J. Immunol.* **148**, 3042–3048 (1992).
10. Monestier, M. & Novick, K. E. Specificities and genetic characteristics of nucleosome-reactive antibodies from autoimmune mice. *Mol. Immunol.* **33**, 89–99 (1996).
11. Carter, R. H., Spycher, M. O., Ng, Y. C., Hoffman, R. & Fearon, D. T. Synergistic interaction between complement receptor type 2 and membrane IgM on B lymphocytes. *J. Immunol.* **141**, 457–463 (1988).
12. Ahearn, J. M. et al. Disruption of the Cr2 locus results in a reduction in B-1a cells and an impaired B cell response to T-dependent antigen. *Immunity* **4**, 251–262 (1996).
13. Lemaître, B., Nicholas, E., Michaut, L., Reichhart, J. M. & Hoffmann, J. A. The dorsoventral regulatory gene cassette *spatzle/Toll/cactus* controls the potent antifungal response in *Drosophila* adults. *Cell* **86**, 973–983 (1996).
14. Medzhitov, R., Preston-Hurlburt, P. & Janeway, C. A. A human homologue of the *Drosophila* Toll protein signals activation of adaptive autoimmunity. *Nature* **388**, 323–324 (1997).
15. Akira, S., Takeda, K. & Kaisho, T. Toll-like receptors: critical proteins linking innate and acquired immunity. *Nature Immunol.* **2**, 675–680 (2001).
16. Li, M. et al. An essential role of the NF- κ B/Toll-like receptor pathway in induction of inflammatory and tissue-repair gene expression by necrotic cells. *J. Immunol.* **166**, 7128–7135 (2001).
17. Horng, T., Barton, G. M. & Medzhitov, R. TIRAP: an adapter molecule in the Toll signalling pathway. *Nature Immunol.* **2**, 835–841 (2001).
18. Adachi, O. et al. Targeted disruption of the *Myd88* gene results in loss of IL-1 and IL-18-mediated function. *Immunity* **9**, 143–150 (1998).
19. Hacker, H. et al. Immune cell activation by bacterial CpG-DNA through myeloid differentiation marker 88 and tumour necrosis factor receptor-associated factor (TRAF)6. *J. Exp. Med.* **192**, 595–600 (2000).
20. Singal, R. & Ginder, G. D. DNA methylation. *Blood* **93**, 4059–4070 (1999).
21. Hemmi, H. et al. A Toll-like receptor recognizes bacterial DNA. *Nature* **408**, 740–745 (2000).
22. Yi, A.-K. et al. CpG motifs in bacterial DNA activate leukocytes through the pH-dependent generation of reactive oxygen species. *J. Immunol.* **160**, 4755–4761 (1998).
23. Hacker, H. et al. CpG-DNA-specific activation of antigen-presenting cells requires stress kinase activity and is preceded by non-specific endocytosis and endosomal maturation. *EMBO J.* **17**, 6230–6240 (1998).
24. Benaroch, P. et al. How MHC class II molecules reach the endocytic pathway. *EMBO J.* **14**, 37–49 (1995).
25. Massari, P. et al. Immune stimulation by neisserial porins is TLR2 and MyD88-dependent. *J. Immunol.*

- 168, 1533–1537 (2002).
26. The Canadian Hydroxychloroquine Study Group. A randomized study of the effect of withdrawing hydroxychloroquine sulfate in systemic lupus erythematosus. *N. Engl. J. Med.* **324**, 150–154 (1991).
27. Furst, d. E. *et al.* Dose-loading with hydroxychloroquine improves the rate of response in early, active rheumatoid arthritis. *Arthritis Rheum.* **42**, 357–365 (1999).
28. Lenart, P., Stunz, L., Yi, A.-K., Krieg, A. M. & Ashman, R. F. CpG stimulation of primary mouse B cells is blocked by inhibitory oligodeoxyribonucleotides at a site proximal to NF- κ B activation. *Antisense Nucleic Acid Drug Dev.* **4**, 247–256 (2001).
29. Shlomchik, M. J., Marshak-Rothstein, A., Wolfowicz, C. B., Rothstein, T. L. & Weigert, M. G. The role of clonal selection and somatic mutation in autoimmunity. *Nature* **328**, 805–811 (1987).
30. Zeng, D., Lee, M.-K., Tung, J., Brendolan, A. & Strober, S. Cutting edge: A role for CD1 in the pathogenesis of Lupus in NZB/NZW mice. *J. Immunol.* **164**, 5000–5004 (2000).
31. William, J., Christensen, C. & Shlomchik, M. J. Evolution of an autoantibody response is linked to somatic hypermutation outside of germinal centers. (submitted).
32. Botto, M. *et al.* Homozygous C1q deficiency causes glomerulonephritis associated with multiple apoptotic bodies. *Nature Genet.* **19**, 56–59 (1998).
33. Bickerstaff, M. C. M. *et al.* Serum amyloid P component controls chromatin degradation and prevents antinuclear autoimmunity. *Nature Med.* **5**, 694–697 (1999).
34. Napipei, M. *et al.* Features of systemic lupus erythematosus in DNase1-deficient mice. *Nature Genet.* **25**, 177–180 (2000).
35. Scott, R. S. *et al.* Phagocytosis and clearance of apoptotic cells is mediated by MER. *Nature* **411**, 207–211 (2001).
36. Bell, D. A., Morrison, B. & VandenBygaart, P. Immunogenic DNA-related factors. Nucleosomes spontaneously released from normal murine lymphoid cells stimulate proliferation and immunoglobulin synthesis of normal mouse lymphocytes. *J. Clin. Invest.* **85**, 1487–1496 (1990).
37. Bell, D. A. & Morrison, B. The spontaneous apoptotic death of normal human lymphocytes *in vitro*: the release of, and immunoproliferative response to, nucleosomes *in vitro*. *Clin. Immunol. Immunopathol.* **60**, 1326 (1991).
38. Moller, G., Andersson, J. & Sjoberg, O. Lipopolysaccharides can convert heterologous red cells into thymus-independent antigens. *Cell. Immunol.* **4**, 416–424 (1972).
39. Kawai, T., Adachi, O., Ogawa, T., Takeda, K. & Akira, S. Unresponsiveness of MyD88-deficient mice to endotoxin. *Immunity* **11**, 115–122 (1999).

Acknowledgements

We thank M. Carroll, S. Akira and D. Golenbock for providing the Cr2-deficient and MyD88-deficient mice; M. Boulé, C. Chi, C. Lau and G. Yospin for technical assistance; L. Wetzler and D. Golenbock for providing TLR ligands; H. Ploegh for providing concanamycin B; M. Fenton, L. Wetzler, R. Corley, R. Medzhitov, T. Rothstein and D. Stollar for reviewing the manuscript and/or discussions. This work was supported by grants from the National Institutes of Health, the Arthritis Foundation, and the National Kidney Foundation.

Competing interests statement

The authors declare competing financial interests: see the website (<http://www.nature.com>) for details.

Correspondence and requests for materials should be addressed to A.M.-R. (e-mail: amrothst@bu.edu).

Scaling of entanglement close to a quantum phase transition

A. Osterloh^{*†}, Luigi Amico^{*†}, G. Falci^{*†} & Rosario Fazio^{†‡}

^{*} Dipartimento di Metodologie Fisiche e Chimiche (DMFCI), viale A. Doria 6, 95125 Catania, Italy

[†] NEST-INFM, Piazza dei Cavalieri 7, I-56126 Pisa, Italy

[‡] Scuola Normale Superiore, Piazza dei Cavalieri 7, I-56126 Pisa, Italy

Classical phase transitions occur when a physical system reaches a state below a critical temperature characterized by macroscopic order¹. Quantum phase transitions occur at absolute zero; they are induced by the change of an external parameter or coupling constant², and are driven by quantum fluctuations. Examples include transitions in quantum Hall systems³, localization in Si-MOSFETs (metal oxide silicon field-effect transistors; ref. 4) and the superconductor–insulator transition in two-dimensional systems^{5,6}. Both classical and quantum critical points are governed by a diverging correlation length, although quantum systems possess additional correlations that do not have a classical counterpart. This phenomenon, known as entanglement, is the resource that enables quantum computation and communication⁸. The role of entanglement at a phase transition is not captured by statistical mechanics—a complete classification of the critical many-body state requires the introduction of concepts from quantum information theory⁹. Here we connect the theory of critical phenomena with quantum information by exploring the entangling resources of a system close to its quantum critical point. We demonstrate, for a class of one-dimensional magnetic systems, that entanglement shows scaling behaviour in the vicinity of the transition point.

There are various questions that emerge in the study of this

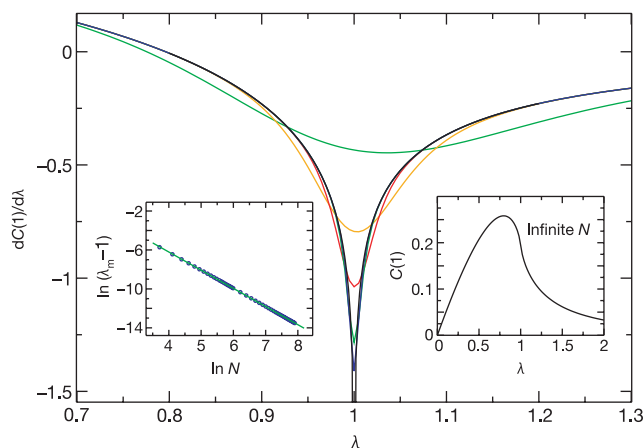


Figure 1 The change in the ground-state wavefunction in the critical region is analysed considering $dC(1)/d\lambda$ as a function of the reduced coupling strength λ . The curves correspond to different lattice sizes $N = 11, 41, 101, 251, 401, \infty$. We choose N odd to avoid the subtleties connected with boundary terms¹⁶. On increasing the system size, the minimum gets more pronounced. Also the position of the minimum changes and tends as $N^{-1.87}$ (left inset) towards the critical point $\lambda_c = 1$ where for an infinite system a logarithmic divergence is present (see equation (3)). The right inset shows the behaviour of the concurrence $C(1)$ itself for an infinite system. The maximum that occurs below λ_c is not related to the critical properties of the Ising model. As explained in the text, it is the change in the ground state and not the wavefunction itself that is a good indicator of the transition. The structure of the reduced density matrix, necessary to calculate the concurrence, follows from the symmetry properties of the hamiltonian. Reality and parity conservation of H together with translational invariance already fix the structure of ρ to be real symmetric with $\rho_{11}, \rho_{22} = \rho_{33}, \rho_{23}, \rho_{14}, \rho_{44}$ as the only non-zero entries.

problem. Because the ground-state wavefunction undergoes qualitative changes at a quantum phase transition, it is important to understand how its genuine quantum aspects evolve throughout the transition. Will entanglement between distant subsystems be extended over macroscopic regions, as correlations are? Will it carry distinct features of the transition itself and show scaling behaviour? Answering these questions is important for a deeper understanding of quantum phase transitions, and also from the perspective of quantum information theory. So results that bridge these two areas of research are of great relevance.

We study a set of localized spins coupled through exchange interaction and subjected to an external magnetic field (we consider only spin-1/2 particles), a model central both to condensed-matter and information theory and subject to intense study¹⁰. In the Heisenberg chain, the maximization of the entanglement at zero temperature is related to the energy minimization¹¹. It is known that Werner states¹² can be generated in a one dimensional XY model¹³, and that temperature and magnetic field can increase the entanglement of the systems, as shown for the Ising and Heisenberg models^{14,15}. Finally, we mention the study of the role of the entanglement in the density matrix renormalization group flow and the introduction of entanglement-preserving renormalization schemes¹⁶. Here we address the problem of the relation between macroscopic order, classical correlations and quantum correlations. Therefore we analyse the entanglement near the critical point of the XY model in a transverse magnetic field. Because of the universality principle—the critical behaviour depends only on the dimension of the system and the symmetry of the order parameter—our results have much broader validity. We find that in the vicinity of a quantum phase transition the entanglement obeys scaling behaviour. On the other hand, this analysis provides a clear distinction between the role of entanglement and correlations in quantum systems close to a critical point. (We have been made aware that similar work to that reported here is being performed by T. Osborne and M. Nielsen; T. Osborne, personal communication.)

The system under consideration is a spin-1/2 ferromagnetic chain with an exchange coupling J in a transverse magnetic field of

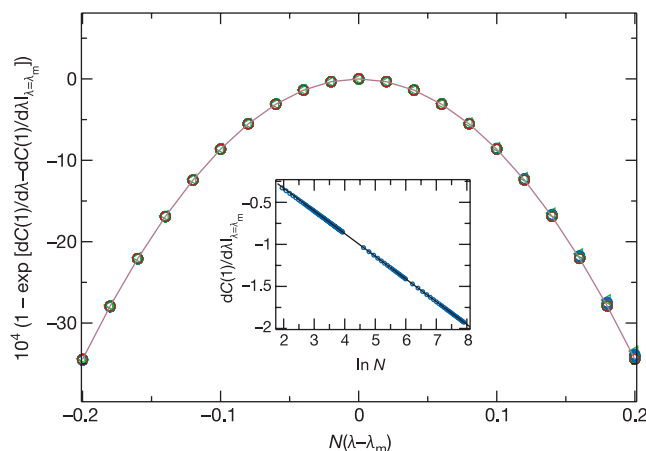


Figure 2 The finite size scaling is performed for the case of logarithmic divergences²². The concurrence, considered as a function of the system size and the coupling constant is a function of $N^{1/\nu}(\lambda - \lambda_m)$ only, and in the case of logarithmic divergence it behaves as $[dC(1)/d\lambda]_{N\lambda} - [dC(1)/d\lambda]_{N\lambda_0} \sim Q[N^{1/\nu}(\lambda - \lambda_m)] - Q[N^{1/\nu}(\lambda_0 - \lambda_m)]$ where λ_0 is a non-critical value and $Q(x) \sim Q(\infty) \ln x$ (for large x). All the data from $N = 41$ up to $N = 2,701$ collapse on a single curve. The critical exponent is $\nu = 1$, as expected for the Ising model. The inset shows the divergence of the value at the minimum as the system size increases.

strength h . The hamiltonian is

$$H = -\frac{J}{2}(1 + \gamma) \sum_{i=1}^N \sigma_i^x \sigma_{i+1}^x - \frac{J}{2}(1 - \gamma) \sum_{i=1}^N \sigma_i^y \sigma_{i+1}^y - h \sum_{i=1}^N \sigma_i^z \quad (1)$$

where σ^a are the Pauli matrices ($a = x, y, z$) and N is the number of sites. We assume periodic boundary conditions. It is convenient, for later purposes, to define a dimensionless coupling constant $\lambda = J/2h$. For $\gamma = 1$ equation (1) reduces to the Ising model, whereas for $\gamma = 0$ it is the XY model. For all the interval $0 < \gamma \leq 1$ the models belong to the Ising universality class and for $N = \infty$ they undergo a quantum phase transition at the critical value $\lambda_c = 1$. The magnetization $\langle \sigma^x \rangle$ is different from zero for $\lambda > 1$ and it vanishes at the transition. On the contrary the magnetization along the z direction $\langle \sigma^z \rangle$ is different from zero for any value of λ . At the phase transition the correlation length ξ diverges as $\xi \sim |\lambda - \lambda_c|^{-\nu}$ with $\nu = 1$ (refs 17 and 18).

We confine our interest to the entanglement between two spins, of position i and j , in the chain. All the information needed is contained in the reduced density matrix $\rho(i, j)$ obtained from the ground-state wavefunction after all the spins except those at positions i and j have been traced out. The resulting $\rho(i, j)$ represents a mixed state of a bipartite system; a good deal of work has been devoted to quantifying the entanglement in this case^{19–22}. We use the concurrence²² between sites i and j , related to the “entanglement of formation”²¹, and defined as

$$C(i, j) = \max\{r_1(i, j) - r_2(i, j) - r_3(i, j) - r_4(i, j), 0\} \quad (2)$$

In equation (2), $r_\alpha(i, j)$ are the square roots of the eigenvalues of the product matrix $R = \rho(i, j)\tilde{\rho}(i, j)$ in descending order; the spin flipped matrix is defined as $\tilde{\rho} = \sigma^y \otimes \sigma^y \rho^* \otimes \sigma^y$. In the previous definition, the eigenstates of $\sigma^z \{|\uparrow\rangle, |\downarrow\rangle\}$ should be used. Translation invariance implies that $C(i, j) = C(|i - j|)$. The concurrence will be evaluated as a function of the relative position $|i - j|$ between the spins and the distance $|\lambda - \lambda_c|$ from the critical point. The structure of the reduced density matrix is obtained by exploiting symmetries of the model (see Fig. 1 legend). The non-zero entries of ρ can then be related to the various correlation functions, and the concurrence of the ground state is evaluated exactly starting from the results in refs 17, 18 and 23.

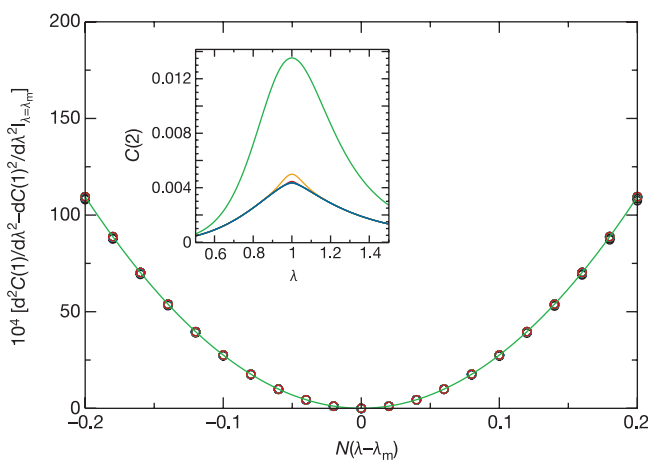


Figure 3 As in the case of the nearest-neighbour concurrence, data collapse is also obtained for the next-nearest-neighbour concurrence $C(2)$. In the figure, data for system size from $N = 41$ to $N = 401$ are plotted. The inset shows a peculiarity of the Ising model: $C(2)$ has its maximum at the critical point for arbitrary system size (note that the maximum decreases as the size increases). Therefore we consider the second derivative to perform the scaling analysis. It can also be seen that $C(2)$ is two orders of magnitude smaller than $C(1)$. For the smallest system sizes, the concurrence is different from zero for $|i - j| = 3$ and $\lambda > 1.05$ (for $N = 7$; for $N \approx 9$ $C(3) = 0$ for all λ). In contrast, the correlation functions are long-ranged at the critical point.

First, we look at the Ising model ($\gamma = 1$). The first question we consider is the range of the entanglement ξ_E , that is, the maximum distance between two spins at which the concurrence is different from zero. The result is surprising: even at the critical point, where spin–spin correlations extend over a long range (the correlation length is diverging for an infinite system), the concurrence vanishes unless the two sites are at most next-nearest neighbours. The truly non-local quantum part of the two-point correlations is nonetheless very short-ranged.

In order to quantify the change of the many-body wavefunction when the system crosses the critical point, we look at the derivatives of the concurrence as a function of λ . In this case we need to consider only the nearest-neighbour and next-nearest-neighbour concurrence. We first discuss the behaviour of the nearest-neighbour concurrence. The results for systems of different size (including the thermodynamic limit) are presented in Fig. 1. For the infinite chain $\partial_\lambda C(1)$ diverges on approaching the critical value as:

$$\frac{dC(1)}{d\lambda} = \frac{8}{3\pi^2} \ln|\lambda - \lambda_c| + \text{const.} \quad (3)$$

Equation (3) quantifies non-local correlations in the critical region. One aspect of this system, particularly relevant for quantum information, is the study of the precursors of the critical behaviour in finite samples. This study is known as finite size scaling²⁴. In Fig. 1 the derivative of $C(1)$ with respect to λ is considered for different system sizes. As expected, there is no divergence for finite N , but there are clear anomalies. The position of the minimum λ_m scales as $\lambda_m \sim \lambda_c + N^{-1.86}$ and its value diverges logarithmically with increasing system size as:

$$\left. \frac{dC(1)}{d\lambda} \right|_{\lambda_m} = -0.2702 \ln N + \text{const.} \quad (4)$$

According to the scaling ansatz²⁴ in the case of logarithmic singularities, the ratio between the two prefactors of the logarithm in equations (3) and (4) is the exponent that governs the divergence of the correlation length ν . In this case ($8/3\pi^2 \approx 0.2702$) it follows that $\nu = 1$, as it is known from the solution of the Ising model¹⁷. By proper scaling²⁴ and taking into account the distance of the minimum of $C(1)$ from the critical point, it is possible to make all the data from different N collapse onto a single curve (Fig. 2). This figure contains the data for the lattice size ranging from $N = 41$ up to $N = 2,701$. These results show that all the key ingredients of the finite size scaling are present in the concurrence. We note that

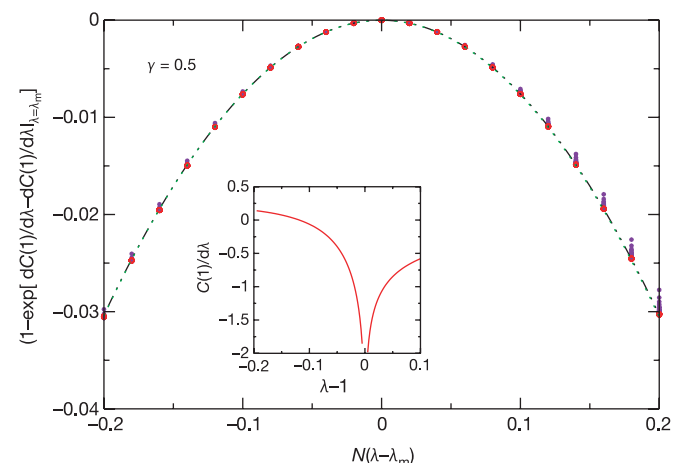


Figure 4 The universality hypothesis for the entanglement is checked by considering the model hamiltonian, defined in equation (1), for a different value of γ . In this case we chose $\gamma = 0.5$ and N ranging from 41 up to 401. Data collapse, shown here for $C(1)$, is obtained for $\nu = 1$, consistent with the model being in the universality class of the Ising model. In the inset is shown the divergence at the critical point for the infinite system.

finite size scaling is fulfilled over a very broad range of values of N which are of interest in several protocols in quantum information.

A similar analysis can be carried on for the next-nearest-neighbour concurrence $C(2)$. Since $\partial_\lambda C(2)|_{\lambda_c} = 0$, the logarithmic singularity here appears first in the second derivative with respect to λ (see Fig. 3 legend). In the thermodynamic limit $\partial_\lambda^2 C(2) = 0.108 \ln|\lambda - \lambda_c| + \text{const.}$ In this case also, the data collapse and finite size scaling (Fig. 3) agree with the expected scaling behaviour, $\nu = 1$. This completes the analysis of entanglement for the one-dimensional Ising model.

A cornerstone of the theory of critical phenomena is the concept of universality—that is, the critical properties depend only on the dimensionality of the system and the broken symmetry in the ordered phase. Universality in the critical properties of entanglement was verified by considering the properties of the family of models defined in equation (1) with $\gamma \neq 1$.

Second, we consider the case for $\gamma \neq 1$. The range of entanglement ξ_E is not universal. The maximum possible distance between entangled pairs increases and tends to infinity as γ tends to zero. From the asymptotic behaviour of the reduced density matrix¹⁸ we find that ξ_E goes as γ^{-1} . This however has no dramatic consequences; the “total concurrence” $\Sigma_n C(n)$ stored in the chain is an increasing function of γ (for $0 < \gamma \leq 1$, $0 < \Sigma_n C(n) < 0.2$). More interesting is the critical behaviour of the concurrence. To be specific we consider $C(1)$ in the case $\gamma = 0.5$, shown in Fig. 4. As it was obtained for the hamiltonian of equation (1), scaling is fulfilled with the critical exponent $\nu = 1$ in agreement with the universality hypothesis.

The analysis of the ‘resource’ entanglement for a condensed-matter system close to a quantum critical point allows us to characterize both quantitatively and qualitatively the change in the wavefunction of the ground state on passing the phase transition. A notable feature which emerges is that though the entanglement itself is not an indicator of the phase transition, an intimate connection exists between entanglement, scaling and universality. In a way, this analysis allows us to discern what is genuinely quantum in a zero-temperature phase transition. The results presented here might be tested by measuring different correlation functions, for example with neutron scattering, and extracting from these the entanglement properties of the ground state close to the critical point¹⁴.

We finally discuss this work in the context of quantum computation. First, system sizes considered here ($\sim 10^3$) could be those of a realistic quantum computer. Second, the scaling behaviour found could be a powerful tool to evaluate (and hence to use) entanglement in systems having different numbers of qubits. In particular, close to the critical point, the entanglement depends strongly on the field—so it could be tuned, realizing an ‘entanglement switch’. Last, long-range correlations, typical of the critical region, might be of great importance in stabilizing the system against errors due to imperfections. □

Received 14 December 2001; accepted 21 February 2002.

1. Goldenfeld, N. *Lectures on Phase Transitions and the Renormalization Group* (Addison Wesley, New York, 1992).
2. Sachdev, S. *Quantum Phase Transitions* (Cambridge Univ. Press, Cambridge, 2000).
3. Sondhi, S. I., Girvin, S. M., Carini, J. P. & Shahar, D. Continuous phase transitions. *Rev. Mod. Phys.* **69**, 315–333 (1997).
4. Kravchenko, S. V., Kravchenko, G. V., Furneaux, J. E., Pudalov, V. M. & D'Iorio, M. Possible metal-insulator transition at $B = 0$ in two dimensions. *Phys. Rev. B* **50**, 8039–8042 (1994).
5. Haviland, D. B., Liu, Y. & Goldman, A. M. Onset of superconductivity in the two-dimensional limit. *Phys. Rev. Lett.* **62**, 2180–2183 (1989).
6. van der Zant, H. S. J., Fritschy, F. C., Elion, W. E., Geerligs, L. J. & Mooij, J. E. Field-induced superconductor-to-insulator transition in Josephson junction arrays. *Phys. Rev. Lett.* **69**, 2971–2974 (1992).
7. Bell, J. S. *Speakable and Unsayable in Quantum Mechanics* (Cambridge Univ. Press, Cambridge, 1987).
8. Nielsen, M. & Chuang, I. *Quantum Computation and Quantum Communication* (Cambridge Univ. Press, Cambridge, 2000).
9. Preskill, J. Quantum information and physics: some future directions. *J. Mod. Opt.* **47**, 127–137 (2000).
10. Brooke, J., Bitko, D., Rosenbaum, T. F. & Aeppli, G. Quantum annealing of a disordered spin system. *Science* **284**, 779–781 (1999).

11. O'Connors, K. M. & Wootters, W. K. Entangled rings. *Phys. Rev. A* **63**, 052302–1–052302–9 (2001).
12. Dür, W., Vidal, G. & Cirac, J. I. Three qubits can be entangled in two inequivalent ways. *Phys. Rev. A* **62**, 062314–1–062314–12 (2000).
13. Wang, X. Entanglement in the quantum Heisenberg XY model. *Phys. Rev. A* **64**, 012313–1–012313–7 (2001).
14. Arnesen, M. C., Bose, S. & Vedral, V. Natural thermal and magnetic entanglement in the 1D Heisenberg model. *Phys. Rev. Lett.* **87**, 017901–1–017901–4 (2001).
15. Gunlycke, D., Bose, S., Kendon, V. M. & Vedral, V. Thermal concurrence mixing in a 1D Ising model. *Phys. Rev. A* **64**, 042302–1–042302–7 (2001).
16. Osborne, T. J. & Nielsen, M. A. Entanglement, quantum phase transitions, and density matrix renormalization. Preprint quant-ph/0109024 at (<http://xxx.lanl.gov>) (2001).
17. Pfeuty, P. The one-dimensional Ising model with a transverse field. *Ann. Phys.* **57**, 79–90 (1970).
18. Barouch, E. & McCoy, B. M. Statistical mechanics of the XY model. II Spin-correlation functions. *Phys. Rev. A* **3**, 786–804 (1971).
19. Bennett, C. J., Bernstein, H. J., Popescu, S. & Schumacher, B. Concentrating partial entanglement by local operations. *Phys. Rev. A* **53**, 2046–2052 (1996).
20. Vedral, V., Plenio, M. B., Rippin, M. A. & Knight, P. Quantifying entanglement. *Phys. Rev. Lett.* **78**, 2275–2279 (1997).
21. Bennett, C. H., DiVincenzo, D. P., Smolin, J. & Wootters, W. K. Mixed-state entanglement and quantum error correction. *Phys. Rev. A* **54**, 3824–3851 (1996).
22. Wootters, W. K. Entanglement of formation of an arbitrary state of two qubits. *Phys. Rev. Lett.* **80**, 2245–2248 (1998).
23. Lieb, E., Schultz, T. & Mattis, D. Two soluble models in antiferromagnetic chain. *Ann. Phys.* **60**, 407–466 (1961).
24. Barber, M. N. in *Phase Transitions and Critical Phenomena* Vol. 8 (eds Domb, C. & Leibovitz, J. L.) 146–259 (Academic, London, 1983).

Acknowledgements

We thank G.M. Palma, F. Plastina and J. Siewert for discussions. This work was supported by the European Community (IST-SQUBIT) and by INFN-PRA-SSQI.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to R.F. (e-mail: fazio@sns.it).

Spontaneous breaking of time-reversal symmetry in the pseudogap state of a high- T_c superconductor

A. Kaminski*†, S. Rosenkranz*†, H. M. Fretwell‡, J. C. Campuzano*†, Z. Li§, H. Raffy§, W. G. Cullen†, H. You†, C. G. Olson||, C. M. Varma¶ & H. Höchst#

* Department of Physics, University of Illinois at Chicago, Chicago, Illinois 60607, USA

† Materials Science Division, Argonne National Laboratory, Argonne, Illinois 60439, USA

‡ Department of Physics, University of Wales Swansea, Swansea SA2 8PP, UK

§ Laboratoire de Physique des Solides, Université Paris-Sud, 91405 Orsay, France

|| Ames Laboratory, Iowa State University, Ames, Iowa 50011, USA

¶ Bell Laboratories, Lucent Technologies, Murray Hill, New Jersey 07974, USA

Synchrotron Radiation Center, Stoughton, Wisconsin 53589, USA

A change in ‘symmetry’ is often observed when matter undergoes a phase transition—the symmetry is said to be spontaneously broken. The transition made by underdoped high-transition-temperature (high- T_c) superconductors is unusual, in that it is not a mean-field transition as seen in other superconductors. Rather, there is a region in the phase diagram above the superconducting transition temperature T_c (where phase coherence and superconductivity begin) but below a characteristic temperature T^* where a ‘pseudogap’ appears in the spectrum of electronic excitations^{1,2}. It is therefore important to establish if

T^* is just a cross-over temperature arising from fluctuations in the order parameter that will establish superconductivity at T_c (refs 3, 4), or if it marks a phase transition where symmetry is spontaneously broken^{5–10}. Here we report that, for a material in the pseudogap state, left-circularly polarized photons give a different photocurrent from right-circularly polarized photons. This shows that time-reversal symmetry is spontaneously broken¹¹ below T^* , which therefore corresponds to a phase transition.

We have investigated the time-reversal invariance of the electronic states of $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+\delta}$ (Bi-2212) by ARPES (angle-resolved photoelectron spectroscopy), which becomes sensitive to this symmetry by the use of circularly polarized photons¹¹. The measured ARPES intensity I_α is proportional to $|M_\alpha|^2$, where the matrix element $M_\alpha = \langle \mathbf{p} | O_\alpha | \psi(\mathbf{k}) \rangle$ describes the ejection of an electron from an initial state $|\psi(\mathbf{k})\rangle$ to a final state $|\mathbf{p}\rangle$, and the dipole operator O_α contains the vector potential of $\alpha = \text{L}$ (left) or $\alpha = \text{R}$ (right) circularly polarized photons. The experimental set-up is shown in Fig. 1. It consists of a plane grating monochromator beamline at the Synchrotron Radiation Center in Wisconsin as a source of linearly polarized photons, quadruple reflection polarizer¹², refocusing mirror, and the experimental chamber. It is crucial in this experiment, which is essentially measuring absolute intensity changes, to minimize beam movement, as extraneous intensity changes can occur from such movements. We therefore monitor

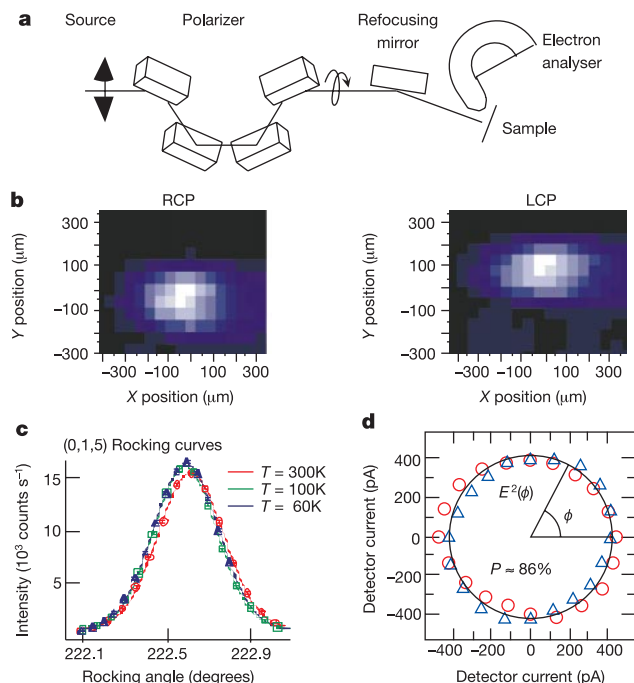


Figure 1 Schematic layout and accuracy of the experimental arrangement. **a**, Experimental set-up. The energy resolution was chosen to be 30 meV. The analyser was operating in angle resolved mode, acquiring 21 energy distribution curves over a range of 5° . The sample positioning precision was $25 \mu\text{m}$. The samples were mounted in normal incidence geometry with a precision of 0.1° using the beamline integrated laser. During the experiment, the angular position of the samples was maintained to within 0.06° using an external laser beam. **b**, Motion of the ultraviolet beam during rotation of the polarizer. **c**, Rocking curves (angular distribution of equivalent perfect crystallites) of the Bi-2212 thin films, showing that there is no distortion of the CuO_2 planes and no rotation of the mirror plane along the Cu–O bonds to within 0.05° . **d**, Degree of circular polarization for right-circularly polarized (RCP; red circles) and left-circularly polarized (LCP; blue triangles) light as used in the experiments. For purely circular polarized light, the value of the electric field vector $E^2(\phi)$ in radial coordinates would be constant (dashed line), whereas for linear polarized light, $E^2(\phi)$ would take the form of a figure eight, being zero perpendicular to the axis of polarization.

the beam position, as shown in Fig. 1, finding a small residual beam movement of $150 \mu\text{m}$, which is compensated for by adjusting the experimental chamber position as the polarizer is rotated. In the experiment we wish to maximize the product TP^2 , where T denotes the transmission and P the polarization. In our experiment, this is accomplished with $P = 86\%$, as shown in Fig. 1d.

Thin-film Bi-2212 samples (1,000–2,000 Å thick, c axis perpendicular to the surface) of various dopings were grown using magnetron sputtering on a SrTiO_3 substrate. The choice of thin-film samples was dictated by the requirement of absolute flatness. Another desirable property of these samples is the small superlattice signal, less than 3% in both ARPES and X-ray diffraction. Using synchrotron X-ray scattering, we have verified that there are no symmetry changes (for example, induced by the substrate) over the temperature range of the measurements (Fig. 1c). The residual magnetization in the experimental chamber was 30 nT. The samples were cleaved *in situ* and measured at pressures $< 3 \times 10^{-11}$ torr.

The symmetry properties of the ARPES matrix elements can be determined with respect to a mirror plane m of the crystal, which is perpendicular to the surface. If we first consider, for simplicity, a time reversal symmetric initial state $|\psi(\mathbf{k})\rangle$, we find two distinct cases: (1) if the vector describing the propagation of the light $\hat{\mathbf{q}}$, the normal to the surface $\hat{\mathbf{n}}$, and the final state momentum $\mathbf{p}$ are not all in m , then there is circular dichroism arising simply from the experimental geometry, owing to the non-coplanarity of the three vectors in question. (2) If all three vectors $\hat{\mathbf{q}}$, $\hat{\mathbf{n}}$ and $\mathbf{p}$ lie in m , then there is no dichroism for time-reversal invariant initial states.

Although case (2) is the one that interests us, we first describe case (1), as this is a large effect, and needs to be correctly accounted for in the analysis of the data^{13–15}. In Fig. 2a we show the geometrical arrangement for case (1) when $\hat{\mathbf{q}}$ and $\hat{\mathbf{n}}$ lie in the mirror plane m

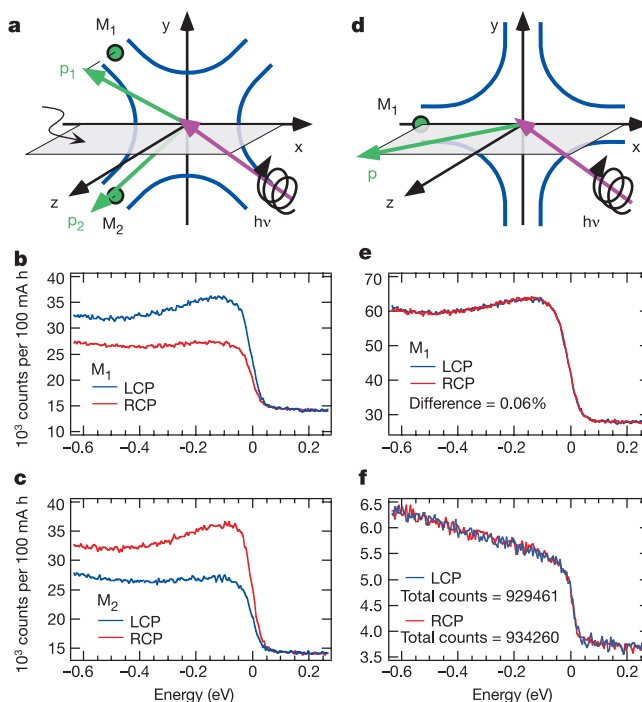


Figure 2 Illustration of the geometric dichroism in ARPES depending on the experimental set-up. **a**, Geometric arrangement of the incident photon, final momentum of the photoemitted electron and symmetry axis when they are not all in the same plane. **b**, Spectra at the M_1 point with LCP and RCP polarizations from an optimally doped sample at $T = 300 \text{ K}$. **c**, Spectra at the M_2 point with LCP and RCP polarizations from the same sample as in **b**. **d**, Geometrical arrangement when the incident photon, final state momentum and symmetry direction are all in the same plane. **e**, Spectra at the M_1 point with LCP and RCP, showing the accuracy of the experiment of $\pm 0.06\%$ from an optimally doped sample at $T = 300 \text{ K}$. **f**, Spectra from polycrystalline Au with LCP and RCP.

diagonal to the Cu–O bond direction, but the final state momenta $\mathbf{p}_1$ at point M_1 , corresponding to $(\pi, 0)$, and $\mathbf{p}_2$ at M_2 , corresponding to $(0, \pi)$, are not in this plane. We can see in Fig. 2b and c that the intensity of the spectra at M_1 for the left-circularly polarized (LCP) light is bigger than that for the right-circularly polarized (RCP) light. This situation is exactly opposite at the point M_2 as expected, because this point is a mirror reflection of the point M_1 about m (ref. 14). On the other hand, when the three vectors $\hat{\mathbf{q}}_p$, $\hat{\mathbf{n}}$ and $\mathbf{p}$ are all in a mirror plane m , as illustrated in Fig. 2d where m is the mirror plane along the Cu–O bond direction, there is no dichroism (although the Cu–O bond direction is strictly speaking not a symmetry direction of the crystal in the presence of a superlattice on the Bi–O plane, it

was found experimentally¹⁶, and confirmed here, that it is in fact a symmetry direction for the CuO_2 planar electronic states). This is seen in Fig. 2e, where we plot spectra obtained at point M_1 that is now in m . In this case, the spectra have equal intensities to an accuracy of $\pm 0.06\%$, which sets the overall accuracy of the experiment. As a test, in Fig. 2f we also show spectra from polycrystalline Au, which also do not exhibit dichroism as the orientation of mirror planes is random.

The interesting situation arises when the initial state $|\psi(\mathbf{k})\rangle$ breaks time-reversal symmetry characterized by an order parameter θ . Therefore for $\mathbf{k}$ in some mirror planes m , $|\psi(\mathbf{k})\rangle$ is not an eigenstate of the reflection operator, even though the charge density retains the same lattice symmetries as for $\theta = 0$. In this case, it has been shown quite generally¹⁵ and for a particular case¹¹ that there is a difference in photoelectron intensity for LCP and RCP light even for $\mathbf{k}$ in m (where the geometrical effect is absent). This effect, proportional to θ , is even for small variations of $\mathbf{k}$ about the mirror plane (while the geometric effect is odd). In order to experimentally detect such a time reversal symmetry breaking (TRSB) state we can plot the energy-integrated intensity¹⁷ (-600 to 100 meV about the chemical potential) of the ARPES signal as a function of the momentum over a very small range ($1/20$ of the Brillouin zone) along a section perpendicular to the mirror plane. In the absence of TRSB, the energy-integrated intensity as a function of momenta will be a straight line, with the slope changing sign for the opposite circular polarization because of the odd symmetry of the geometrical effect. The two lines will cross at the mirror plane where both LCP and RCP intensities are equal. If upon cooling a TRSB state emerges, it will lead to a difference between the RCP and LCP at the symmetry plane.

In Fig. 3a we show the experimental geometry, and Fig. 3b shows the energy-integrated intensity I_L and I_R as a function of momentum along a section indicated in Fig. 3a for an overdoped ($T_c = 64$ K) sample which does not exhibit a pseudogap^{1,2}, as seen from the energy spectra at $\mathbf{k}_f$ in Fig. 3d. In Fig. 3c we plot the dichroism signal $D = (I_R - I_L)/(I_R + I_L)$ obtained from the data shown in Fig. 3b. We see that the dichroism signal is independent of temperature and vanishes at the mirror plane, indicating the absence of TRSB in this sample (The non-zero D away from the mirror plane is just the geometric effect discussed above).

We now perform the same measurements on an underdoped sample with a T_c of 85 K, which has a pseudogap and therefore does exhibit a change of the leading edge position of the photoemission

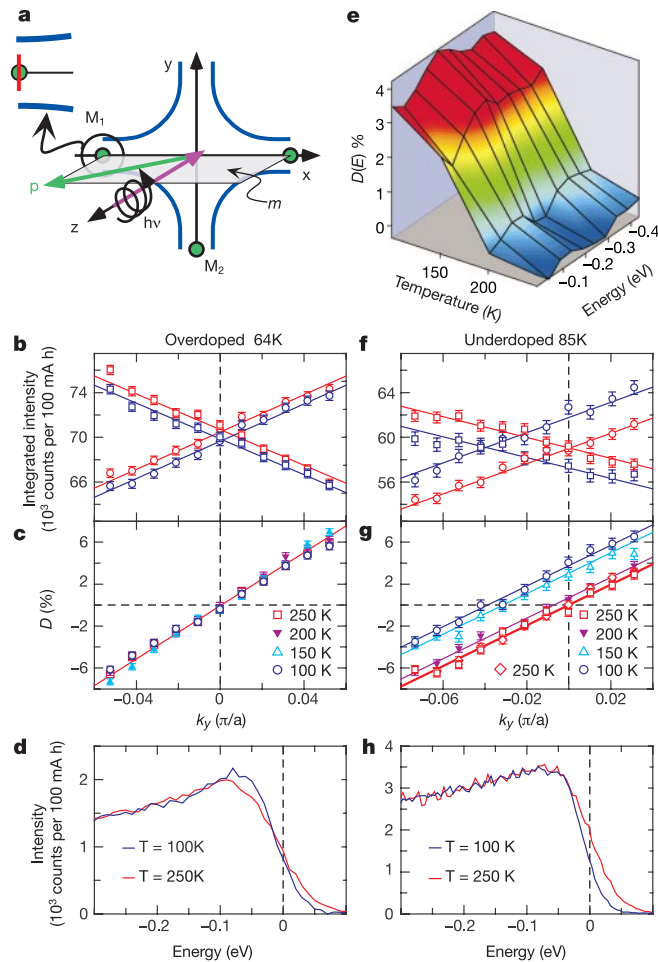


Figure 3 Results of dichroism experiments in overdoped and underdoped samples. **a**, Geometrical arrangement for the experiment, where the incident photon direction, the final state momentum and the normal to the surface are all in the mirror plane. The left side shows an enlarged region around the $(\pi, 0)$ point and the range (red line) covered by the detector. **b**, Energy integrated intensity as a function of momentum perpendicular to the mirror plane at the M_1 point of the BZ for LCP and RCP for an overdoped ($T_c = 64$ K) sample for $T = 250$ K (red) and $T = 100$ K (blue). The lines are linear fits to the data. The data were normalized to the second order. The error bars are statistical only and represent one standard deviation. **c**, Relative difference $D = (I_L - I_R)/(I_L + I_R)$ of the data shown in **b** and in addition for $T = 200$ K (purple) and 150 K (light blue), indicating no circular dichroism for a sample with no shift in the leading edge with temperature. **d**, **e**, Difference signal as a function of temperature and energy, $D(E)$ (integrated over 50 meV). **f**, Energy integrated intensity as a function of momentum perpendicular to the mirror plane at the M_1 point of the Brillouin zone for LCP and RCP for an underdoped sample ($T_c = 85$ K) that displays a shift in the leading edge (**h**), showing no dichroism for $T = 250$ K but clearly showing that there is dichroism for $T = 100$ K. **g**, Relative difference D of the data shown in **f** and in addition for $T = 200$ K and $T = 150$ K. The diamond data points for $T = 250$ K were taken at the end of the experiment after a cooling cycle.

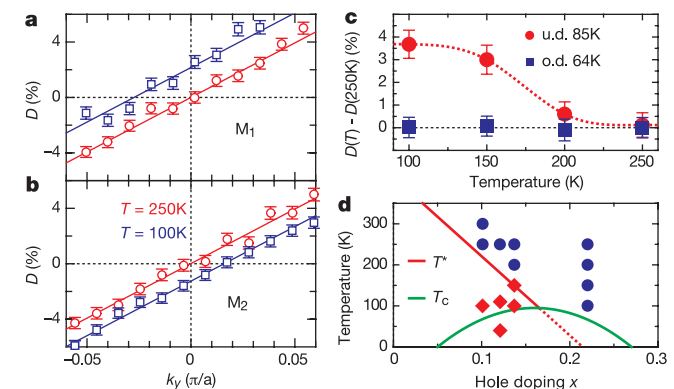


Figure 4 Symmetry and temperature dependence of the observed dichroism. **a**, Relative difference measured for an underdoped sample at M_1 and **b**, after rotating the sample by 90° at M_2 . **c**, Temperature dependence of the dichroism signal at M_1 from the data shown in Fig. 3c (overdoped, o.d., 64 K) and 3f (underdoped, u.d., 85 K). **d**, Phase diagram of Bi-2212 showing the presence (red diamonds) and absence (dark blue dots) of dichroism at the mirror plane along the Cu–O bonds.

spectrum with temperature (Fig. 3h). The momentum dependence of the integrated intensity, Fig. 3f, shows that at high temperatures the LCP and RCP data cross at the mirror plane and therefore there is no TRSB. However, when the sample is cooled to below T^* where the pseudogap opens, we observe a shift of the crossing point of the RCP and LCP data. This shift (2.3°) is much bigger than the temperature-induced structural changes (0.05°), as seen in the rocking curves in Fig. 1, and angular errors in the ARPES experiment (0.06°). We therefore have to conclude that now the intensities for LCP and RCP are not the same at the mirror plane by a difference much bigger than any experimental uncertainty. This is seen more clearly in Fig. 3g, which shows the relative difference D for several temperatures. Each data point in Fig. 3 is the result of several measurements with alternating polarization directions. We further checked that subsequent warming of the sample through T^* gave reproducible results, as shown in Fig. 3g, which rules out sample ageing as the origin of the effect. The constant value of the difference signal with energy (Fig. 3e) indicates that this effect is related to changes of the matrix elements and is not due to artefacts arising from changes in the spectral function¹⁷. This observation of dichroism at the mirror plane provides direct evidence for TRSB in the pseudogap state of Bi-2212.

Additional information about the observed effect can be obtained by examining its symmetry. At $(\pi, 0)$, the changes of D are independent of k along the small section perpendicular to the mirror plane (Fig. 3g) and therefore the observed TRSB effect is even about the mirror plane along the Cu–O bond direction. Furthermore, from the dichroism data obtained from an underdoped sample ($T_c = 78$ K) at two adjacent M points (that is, $M_1 = (\pi, 0)$ and $M_2 = (0, \pi)$) shown in Fig. 4a and b, we conclude that the effect is odd with respect to the mirror plane diagonal to the Cu–O bond direction. These data were obtained by first measuring the spectra at M_1 for $T = 200$ K, then cooling the sample and measuring at $T = 100$ K (Fig. 4a). While keeping the temperature constant, the sample was then rotated by 90° and M_2 was measured at low temperature. After this, the sample was warmed to 200 K and measured again (Fig. 4b). At low temperatures, the RCP signal at M_1 is larger than the LCP signal and therefore D is positive. At M_2 the opposite occurs, and D is negative.

We have performed nine measurement sequences on four different underdoped and two overdoped samples. The sign and absolute value of the dichroism effect is different for different samples, which we attribute to a variation of the number of the two possible domains formed below T^* . Most significantly however, the effect is only observed in underdoped samples below T^* , never above. Our measurements on samples of different doping levels clearly establish that the effect is tied to the pseudogap line (Fig. 4c and d), leading us to conclude that time-reversal symmetry is broken in the pseudogap state of Bi-2212. We also find that, for underdoped samples, the breaking of time-reversal symmetry persists into the superconducting state. □

Received 19 October 2001; accepted 15 February 2002.

- Ding, H. *et al.* Spectroscopic evidence for a pseudogap in the normal state of underdoped high- T_c superconductors. *Nature* **382**, 51–54 (1996).
- Loeser, A. G. *et al.* Excitation gap in the normal state of underdoped $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+\delta}$. *Science* **273**, 325–329 (1996).
- Randeria, M., Trivedi, N., Moreo, A. & Scalettar, R. T. Pairing and spin gap in the normal state of short coherence length superconductors. *Phys. Rev. Lett.* **69**, 2001–2004 (1992).
- Emery, V. J. & Kivelson, S. A. Importance of phase fluctuations in superconductors with small superfluid density. *Nature* **374**, 434–437 (1995).
- Varma, C. M. Non-Fermi-liquid states and pairing instability of a general model of copper oxide metals. *Phys. Rev. B* **55**, 14554–14580 (1997).
- Varma, C. M. Pseudogap phase and the quantum-critical point in copper-oxide metals. *Phys. Rev. Lett.* **83**, 3538–3541 (1999).
- Chakravarty, S., Laughlin, R. B., Morr, D. K. & Nayak, C. Hidden order in the cuprates. *Phys. Rev. B* **63**, 094503–1–094503–10 (2001).
- Tallon, J. L. & Loram, J. W. The doping dependence of T^* —what is the real high- T_c phase diagram? *Physica C* **349**, 53–68 (2001).
- Hsu, T. C., Marston, J. B. & Affleck, I. Two observable features of the staggered-flux phase at nonzero doping. *Phys. Rev. B* **43**, 2866–2877 (1991).

- Ivanov, D. A., Lee, P. A. & Wen, X.-G. Staggered-vorticity correlations in a lightly doped t-J model: A variational approach. *Phys. Rev. Lett.* **84**, 3958–3961 (2000).
- Varma, C. M. Proposal for an experiment to test a theory of high-temperature superconductors. *Phys. Rev. B* **61**, R3804–R3807 (2000).
- Höchst, H., Patel, R. & Middleton, F. Multiple-reflection lambda/4 phase-shifter—a viable alternative to generate circular-polarized synchrotron radiation. *Nucl. Instrum. Methods Phys. Res. A* **347**, 107–114 (1994).
- Dubbs, R. *et al.* Circular dichroism in photoelectron angular distributions from oriented linear molecules. *Phys. Rev. Lett.* **54**, 1249–1251 (1985).
- Venus, D. Magnetic circular dichroism in angular distributions of core-level photoelectrons. *Phys. Rev. B* **48**, 6144–6151 (1993).
- Simon, M. E. & Varma, C. M. Detection of time-reversal breaking state in normal state of cuprate metals. Preprint cond-mat/0201036 at (<http://xxx.lanl.gov>) (2002).
- Ding, H. *et al.* Electronic excitations in $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_8$: Fermi surface, dispersion, and absence of bilayer splitting. *Phys. Rev. Lett.* **76**, 1533–1536 (1996).
- Randeria, M. *et al.* Momentum distribution sum rule for angle-resolved photoemission. *Phys. Rev. Lett.* **74**, 4951–4954 (1995).

Acknowledgements

We thank M. Randeria for discussions, and T. Kubala, G. Rogers and M. Bissen for their help with the polarizer. This work was supported by the US NSF and the US Department of Energy—Basic Energy Sciences. The Synchrotron Radiation Center is supported by the NSF. S.R. is supported by the Swiss NSF.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to J.C.C. (e-mail: jcc@uic.edu).

Optical studies of solid hydrogen to 320 GPa and evidence for black hydrogen

Paul Loubeyre*, Florent Occelli* & René LeToullec*†

* Département Physique Théorique et Applications, Commissariat à l’Energie Atomique, 91680, Bruyères-le-Châtel, France
† Université Paris 6, PMC, 4 place Jussieu, 75252 Paris, France

The quest for metallic hydrogen at high pressures represents a longstanding problem in condensed matter physics^{1,2}. Recent calculations^{3–6} have predicted that solid hydrogen should become a molecular metal at pressures above 300 GPa, before transforming into an alkali metal; but the strong quantum nature of the problem makes the predictions difficult. Over a decade ago, an optical study⁷ of hydrogen was made using a diamond anvil cell to reach 250 GPa. However, despite many subsequent efforts, quantitative studies^{8–11} at higher pressures have proved difficult and their conclusions controversial. Here we report optical measurements of solid hydrogen up to a pressure of 320 GPa at 100 K. The vibron signature of the H_2 molecule persists to at least 316 GPa; no structural changes are detected above 160 GPa, and solid hydrogen is observed to turn completely opaque at 320 GPa. We measure the absorption edge of hydrogen above 300 GPa, observing features characteristic of a direct electronic bandgap. This is at odds with the most recent theoretical calculations that predict much larger direct transition energies and the closure of an indirect gap^{3–6}. We predict that metal hydrogen should be observed at about 450 GPa when the direct gap closes.

In recent years, a large body of experimental work on dense hydrogen has been carried out both at static compression^{8–11} (with the use of diamond anvil cells) and at dynamic compression (either with gas guns¹² or lasers¹³). Remarkably, electronic conductivity has now been unambiguously reported in the dense fluid phase^{12,13}. The

direct effect, however, of high temperatures (a few 1,000 K in the course of the shock compression) on this insulator–metal fluid phase transition is still incompletely understood.

The search for metal hydrogen, as an interesting quantum solid for condensed matter physics², requires static pressure at least above 300 GPa coupled to a detailed characterization of the sample. Various reasons^{7,8} have been invoked to explain the limit of 250 GPa under static compression of the spectroscopic characterization of the hydrogen sample: the chemical reaction of dense hydrogen with components of the diamond cell environment, including gasket, pressure calibrant or diamond anvil; the line broadening due to the non-hydrostaticity of hydrogen and the onset of strong fluorescence from the diamond anvils that cause a dramatic decrease of the signal-to-noise ratio; and the red-shift of the diamond window absorption edge that obscures spectroscopic measurements.

Four very-high-pressure experiments at around 100 K were devoted to this project. Natural type Ia diamond anvils with different tips and different bevels were used. This allowed us to

vary the sample geometry and the stress-induced optical contributions of the diamond window. The loading of high-purity n-H₂ was performed at room temperature in the rhenium gasket of a membrane diamond anvil cell. In one case, the hydrogen solid was embedded in helium to test non-hydrostatic effects¹⁴. The ruby luminescence gauge was used to measure the pressure in the first three experiments¹⁵, but this signal gets very weak above 200 GPa. So, in the fourth experiment, we used the first-order Raman phonon frequency of the diamond tip at the hydrogen interface to estimate the sample pressure. A typical spectrum in the 300 GPa range and its analysis are given in Fig. 1.

The Raman spectrum of hydrogen could be measured accurately above 300 GPa. As seen in Fig. 1, there was no significant increase of the luminescence background of the diamond anvils to hide the hydrogen spectra. A continuous broadening with pressure of the vibron peak was observed that is more pronounced above 160 GPa in phase III. It is reproducible for the various samples (also with the helium quasi-hydrostatic medium) and reaches 206 cm⁻¹ at 316 GPa. In contrast, the low-frequency libron modes remain

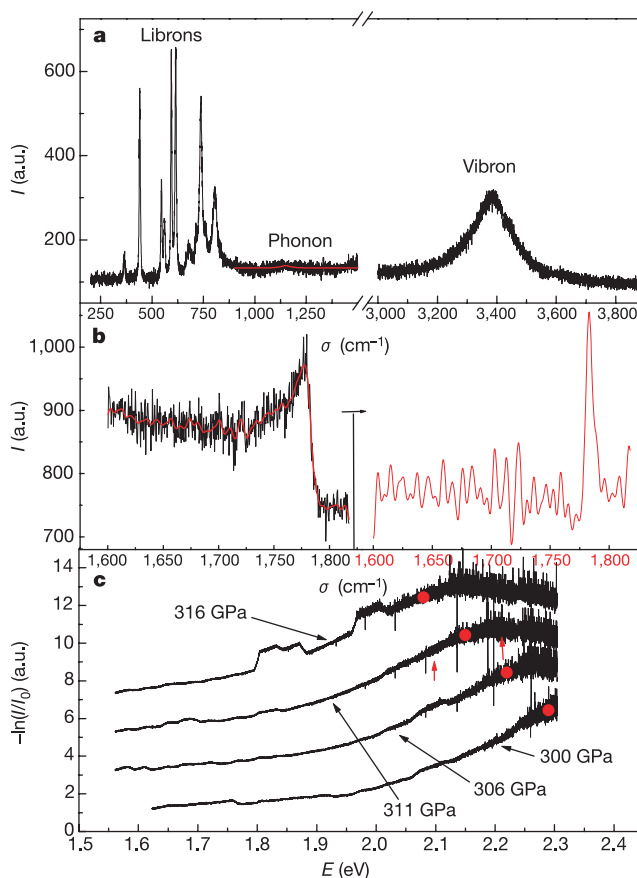


Figure 1 Optical measurements in solid hydrogen at ultra-high pressure. **a**, Raman spectra of the excitations of solid hydrogen at 306 GPa. The spectrum was measured with 30 mW of 647.1-nm radiation for 30 s accumulation time. **b**, Pressure measurement with the Raman spectra of the diamond anvil tip, here at 290 GPa. The spread to high frequency of the degenerated longitudinal transverse optical phonon Raman line of diamond (1,333 cm⁻¹ at ambient pressure) reflects the pressure gradient through the anvil. The high-frequency part that is shown here is coming from the diamond tip (left, in black). The differentiation of this spectra gives a sharp peak (right, in red) that corresponds to the phonon frequency of the diamond tip at the interface with hydrogen. The pressure shift of this peak was calibrated to 200 GPa by using the pressure shift of the vibron frequency of hydrogen (measured reproducibly in three experiments). The diamond-tip phonon frequency was then used as the ultra-high pressure gauge under a Gruneisen expansion form to second order with the equation of state of diamond measured to

150 GPa from single-crystal X-ray diffraction (giving the bulk modulus to be 447 GPa and its pressure derivative to be 3.0). The fitted Gruneisen parameter and its volume derivative were, respectively, 0.9075 and 0.8262. We note that this fit is exactly valid only for the present ultra-high-pressure experiment. The accuracy of the pressure measurement around 300 GPa was estimated to be ± 3 GPa relative to the ruby scale¹⁵. **c**, Absorption edges of solid hydrogen above 300 GPa. Transmission spectra through the hydrogen sample, I , are divided by the transmission of the empty diamond anvil cell, I_0 . The curves at 306 GPa and 316 GPa are vertically translated respectively by 2, 4 and 6 units. The red dots indicate the exciton energy at various pressures, as explained in the text. The two arrows indicate, on the curve at 311 GPa, the change in the positioning of the exciton energy that would correspond to a variation of ± 0.05 eV (estimated error bar). Without translation, the four curves merge above their red dots into a plateau of strong absorption, corresponding to an absorption coefficient of 30,000 cm⁻¹.

sharp up to the maximum pressure. This suggests that the large width of the vibron is an intrinsic property of hydrogen. A detailed analysis of the Raman spectra is beyond the scope of this work, but we can say that the number and the pressure shift of the librational modes, the phonon mode and the vibron mode are consistent with previous measurements in phase III to 220 GPa (refs 9, 10). In particular, as seen in Fig. 2, no discontinuity of the vibron frequency versus pressure is observed above 160 GPa (yet, a temperature shift on the vibron frequency is observed that could be due to orientational molecular ordering¹⁶). The vibrational spectrum is particularly sensitive to changes in structure, so this continuous pressure shift indicates that phase III of molecular hydrogen remains stable up to 316 GPa at least. Also, observation of the vibron mode proves that the molecular form of hydrogen persists to the maximum pressure.

The vibron intensity, σ , versus pressure, p , is also plotted in Fig. 2, the different curves corresponding to the different laser excitation wavelengths. Clear maxima are observed, at higher pressures for

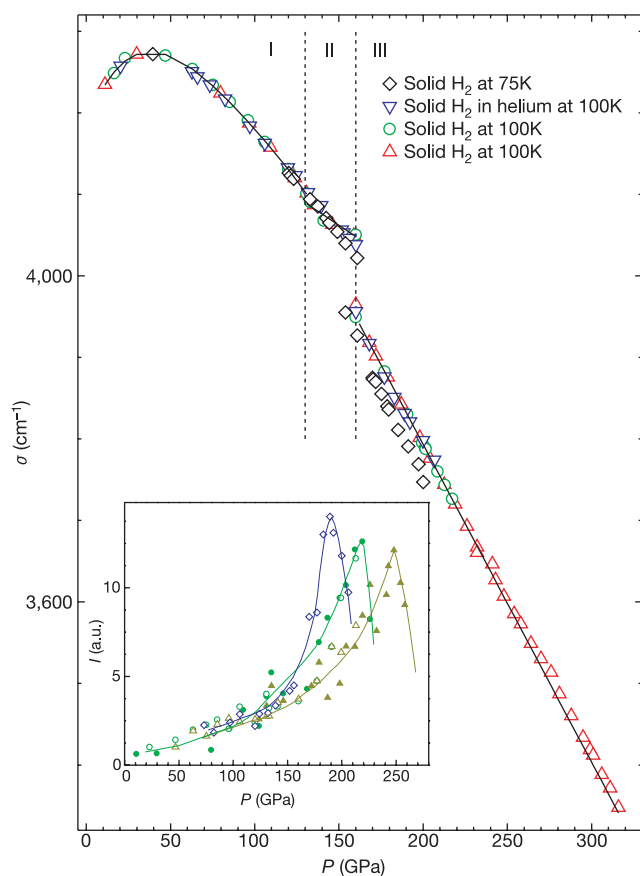


Figure 2 Pressure shift of the Raman H_2 vibron. Different symbols indicate different runs. Experiment 1, on pure H_2 with anvils of 50- μm tip and 7° bevel, achieved a maximum pressure of 200 GPa at 75 K. Experiment 2, on H_2 solid embedded in helium with 50 μm tips and 7° bevel, achieved 206 GPa at 100 K. Experiment 3, on pure hydrogen with anvils of 30- μm tip and 7° bevel, achieved 220 GPa at 100 K. In those runs, the sample diameter was typically of the order of 15 μm diameter. To increase significantly the maximum pressure, a smaller tip (25 μm) a larger bevel (8°) and a very small hydrogen sample (4 μm in diameter) were used in the fourth experiment to achieve 320 GPa. A linear fit reproduces well the data in phase III: $\sigma(\text{in cm}^{-1}) = 4572.8 - 3.9P(\text{in GPa})$. The error bars in pressure (± 3 GPa around 300 GPa and ± 1 GPa around 200 GPa) and error bars in frequency ($\pm 3 \text{ cm}^{-1}$) are smaller than the symbols. Inset, intensity of the vibron peak versus pressure. The different colours correspond to different excitation wavelengths: blue, 488 nm; green, 514.5 nm; dark yellow, 568.2 nm. Open and filled symbols correspond, respectively, to experiments 3 and 4.

longer wavelengths. These curves have been reproducibly observed for various samples and the same magnitude of intensity enhancement has been previously reported⁷ around 180 GPa for the 488-nm curve. The phenomenon is clearly associated with changes in the electronic structure of solid hydrogen. (It is not observed on the intensity of the Raman phonon of the diamond tip.) Resonant Raman intensities should be observed when the wavelengths involved in the Raman process become equal to the electronic levels that build the polarizability of the system. Then, the two peaks of resonance (generally only one, their average, is observed) should be revealed in the scattering efficiency of the sample versus the laser frequency¹⁷. ω_L : the in-going resonance peak when ω_L is equal to the electronic level (here, the exciton) and the out-going resonance peak when ω_L is equal to the electronic level plus the probed excitation of the system (here, the exciton plus the vibron). Therefore, the enhancement of intensity of the H_2 vibron suggests a resonance effect attributable to the exciton of solid hydrogen. We could observe this here at fixed laser wavelengths by varying the pressure because the exciton energy decreases under pressure.

Unfortunately, the precise corrections of absorption necessary to obtain the Raman efficiency of the hydrogen sample are very difficult to achieve. The absorption of the hydrogen sample is unknown in this pressure range and could be measured only above 300 GPa, as explained below. Also, the tail of the absorption edge of the diamond window can strongly perturb the measurement. Both absorption contributions produce the same result, that is, they cut the resonance in the Raman intensity before it reaches the true maximum of the Raman efficiency. Therefore, the pressure

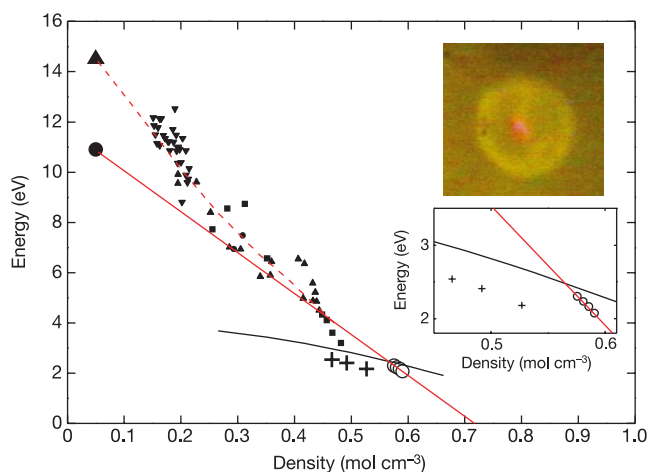


Figure 3 Exciton and direct bandgap of hydrogen as a function of density. Pressure was converted to density with the experimental equation of state¹⁴. The small symbols correspond to the refractive index data: downward triangles, ref. 20; upward triangles, ref. 21; squares, ref. 22. The large filled dot and triangle correspond, respectively, to the ambient pressure exciton and direct gap (ref. 19). Inset, the crosses indicate the determination of the lower bond of the exciton from resonance in the vibron intensity, as explained in the text and the open circles are the exciton data obtained from absorption measurements. The dark solid line is the lower bound for the electronic gap in diamond, indicating the energy below which absorption measurements in a diamond anvil cell have to be carried out at a given pressure¹⁸. The red solid line and red dashed line indicate, respectively, the evolution of the exciton and direct bandgap of hydrogen with density. Inset, the region of the curve around 300 GPa is enlarged. The error bars in pressure (± 3 GPa, corresponding to $\pm 0.003 \text{ mol cm}^{-3}$) and in energy ($\pm 0.05 \text{ eV}$) fall within the circles. The photomicrograph shows reflectance and transmission visual observation of the sample and the tip region (25 μm diameter) is shown at 305 GPa. Incident white light of a quartz-halogen lamp appears red through the hydrogen sample. At 320 GPa, the hydrogen sample appears totally black (less than 2×10^{-3} of the visible white light was going through it).

at the maximum of the vibron intensity for a given excitation ω_L can only give a lower bound for the pressure at which the exciton of solid hydrogen reaches ω_L . Figure 3 shows that these lower bounds for the exciton versus pressure (converted to hydrogen density¹⁴) have the same slope as the shift of the absorption edge of the diamond window¹⁸, thus proving that the absorption coming from the diamond window has indeed strongly perturbed the measurements of the vibron intensity.

An exciton at 10.9 eV and a direct bandgap of 14.5 eV have been determined from absorption edges in solid hydrogen at ambient pressure¹⁹. But the extension of such measurements under pressure has been hampered by the strong absorption of the diamond anvil in the near-ultraviolet spectrum (above 4 eV at ambient pressure). Also, measurements of the absorption edge of type Ia diamond anvil to 405 GPa have shown that the transmission window of diamond closes under high stress¹⁸. Hence, only high-pressure measurements of the index of refraction in the visible spectrum could provide the pressure evolution of an effective oscillator frequency^{20–22} that should be consistent with the evolution of the bandgap. Fortunately, the direct bandgap of hydrogen closes faster under pressure than the bandgap of the diamond window and so they are expected to cross under pressure. Above that crossing point, the absorption edge of solid hydrogen would no longer be obscured by the absorption of the diamond window and consequently information on the electronic state of dense hydrogen could be obtained with confidence. Figure 3 shows that this takes place around 300 GPa. Indeed, the quantitative transmission of solid hydrogen could be measured above 300 GPa, as shown in Fig. 1c. This seems to have the features of a direct absorption edge²³: a steeply rising absorption front to reach a plateau of strong absorption which corresponds to an absorption coefficient of $30,000 \text{ cm}^{-1}$ typical for a direct bandgap (the thickness of hydrogen above 300 GPa was estimated to be of the order of $2 \mu\text{m}$). The exciton edge was located for each pressure by finding the energy at which the rising absorption coefficient sharply bends and becomes flat²⁴ (indicated with red dots in Fig. 1). The low-energy part of the absorption spectra is probably dominated by the tail of the absorption of the diamond anvils under strong stress but could also be partly due to the indirect absorption processes of the hydrogen sample if solid hydrogen has an indirect bandgap, as predicted in most band structure calculations^{3,4,6}. A quantitative analysis of these absorption curves would require knowledge of the structure of phase III, which has not yet been measured.

Figure 3 summarizes the experimental information on the exciton and the direct bandgap of solid hydrogen. Pressures were converted to density with the experimental equation of state¹⁴. The linear fit reproduces well the change of the exciton energy with density from ambient pressure to very high density. This slope is almost independent of the criteria used to locate the exciton and could also be estimated from the shift of the steeply rising part of the absorption edge. This slope is also totally different from the one of the variation of the diamond absorption edge, which makes us more confident that we are examining the absorption of hydrogen here. Indeed, various calculations predict that gap energies should vary almost linearly with density^{3,4}. It should be noted that the measured slope of the density change of the exciton, $-16.3 \text{ eV}/(\text{mol cm}^{-3})$, is smaller than most of the calculated density variation of bandgaps, even for calculations that go beyond the local density approximation^{3,4} (of $-21 \text{ eV}/(\text{mol cm}^{-3})$). The evolution of the direct bandgap is also drawn in Fig. 3 under the following constraints: the effective oscillator energy derived from refractive index data should be slightly greater than the direct bandgap; at high density, excitonic effects are reduced; and the direct bandgap and the exciton should tend to zero at the same density.

These absorption measurements are also clearly translated into the optical aspect of the sample through transmitted white light illumination: at 290 GPa, the hydrogen sample changes from white yellow to orange; under increasing pressure, hydrogen changes to

red; finally, the sample becomes totally opaque to visible light at 320 GPa. From absorption studies of diamond in this pressure range, it was expected that the incident white light should appear yellow in transmission¹⁸. The fact that the absorption of the visible light is due to the hydrogen sample and not to the strained tip of the anvils is also obvious by looking at the sample in reflection. As seen in Fig. 3, the change in colour appears in the sample chamber, whereas the circular area corresponding to the interface between the gasket and the $25 \mu\text{m}$ diamond tip appears yellow and reflective. We note that the pressure in the rhenium gasket at the sample interface is even greater than the pressure in the hydrogen sample by roughly the yield strength of the rhenium gasket⁸. If the darkening of the sample were caused by the absorption of the diamond tip, it would be expected that the surface of the gasket around the sample chamber should also become red and dark in reflection even at lower pressure. Yet, the reflecting metallic gasket and the sample chamber edge could clearly be observed up to the maximum pressure. This formation of black hydrogen is in contrast to ref. 8 where the hydrogen sample was shown to be transparent at 342 GPa. Unfortunately, no measurement of the properties of the hydrogen sample at pressures above 200 GPa was made in ref. 8 that could help to resolve the present discrepancy. We believe that the method of estimating the pressure from the X-ray diffraction of the gasket largely overestimated the pressure given in ref. 8.

From a linear extrapolation in Fig. 3, we see that the direct electronic bandgap should close at 0.71 mol cm^{-3} , which could lead to the formation of atomic metallic hydrogen. The corresponding pressure of about 450 GPa seems to be a more reliable estimate of the metallization pressure than the previous one of 620 GPa that was based on a large extrapolation of the equation of states measured up to 120 GPa (ref. 14). In order to understand the apparent discrepancy between the present observation, showing a direct bandgap closure, and most theoretical predictions^{3–6}, giving indirect bandgap closure and much larger direct absorption energies, the determination of the structure of phase III will be necessary. This work suggests that the quantitative study of solid hydrogen to the 400 GPa range is now within the reach of the diamond anvil cell and that the metallization of hydrogen should therefore be observable at these pressures. $\square$

Received 30 August 2001; accepted 22 February 2002.

- Wigner, E. & Huntington, H. B. On the possibility of a metallic modification of hydrogen. *J. Chem. Phys.* **3**, 764–770 (1935).
- Maksimov, E. G. & Shilov, Y. I. Hydrogen at high pressure. *Phys. Uspekhi* **42**, 1121–1138 (1999).
- Johnson, K. & Ashcroft, N. W. Structure and band gap closure in dense hydrogen. *Nature* **403**, 632–635 (2000).
- Stadler, M. & Martin, R. M. Metallization of molecular hydrogen: prediction from exact-exchange calculations. *Phys. Rev. Lett.* **84**, 6070–6073 (2000).
- Natoli, V., Martin, R. M. & Ceperley, D. Crystal structure of molecular hydrogen at high pressure. *Phys. Rev. Lett.* **74**, 1601–1604 (1995).
- Nagao, K., Nagara, H. & Matsubara, S. Structures of hydrogen at megabar pressures. *Phys. Rev. B* **56**, 2295–2298 (1997).
- Mao, H. K. & Hemley, R. J. Optical studies of hydrogen above 200 GPa: Evidence for metallization by band overlap. *Science* **244**, 1462–1465 (1989).
- Narayana, C., Luo, H., Oroloff, J. & Ruoff, A. L. Solid hydrogen at 342 GPa: no evidence for an alkali metal. *Nature* **393**, 46–49 (1998).
- Goncharov, A. F., Hemley, R. J., Mao, H. K. & Shu, J. New high pressure excitations in parahydrogen. *Phys. Rev. Lett.* **80**, 101–104 (1998).
- Goncharov, A., Gregoryanz, E., Hemley, R. & Mao, H. K. *Proc. Natl Acad. Sci. USA* **98**, 14234–14237 (2001).
- Chen, N., Sterer, E. & Silvera, I. F. Extended infrared studies of high pressure hydrogen. *Phys. Rev. Lett.* **76**, 1663–1666 (1996).
- Weir, S. J., Mitchell, A. C. & Nellis, W. J. Metallization of fluid molecular hydrogen at 140 GPa. *Phys. Rev. Lett.* **76**, 1860–1863 (1996).
- Collins, G. W. *et al.* Measurements of the equation of state of deuterium at the fluid insulator-metal transition. *Science* **281**, 1178–1181 (1998).
- Loubeyre, P. *et al.* X-ray diffraction and equation of state of hydrogen at megabar pressures. *Nature* **383**, 702–704 (1996).
- Mao, H. K., Xu, J. & Bell, P. M. Calibration of the ruby pressure gauge to 800 Kbar under quasihydrostatic conditions. *J. Geophys. Res.* **B 91**, 4673–4676 (1986).
- Goncharov, L. F. *et al.* Invariant points and phase transitions in deuterium at megabar pressures. *Phys. Rev. Lett.* **75**, 2514–2517 (1995).
- Cardona, M. in *Topics in Applied Physics* (eds Cardona, M. & Guntherodt, G.) **Vol. 50**, 19–178 (Springer, New York, 1982).

18. Ruoff, A. L., Luo, H. & Vohra, Y. K. The closing diamond anvil optical window in multimegabar research. *J. Appl. Phys.* **69**, 6413–6416 (1991).
19. Inoue, K., Kanzaki, H. & Suga, S. Fundamental spectra of solid hydrogen. *Solid State Commun.* **30**, 627–629 (1979).
20. Van Straaten, J. & Silvera, I. F. Pressure dependence of the optical absorption edge of solid hydrogen in a diamond anvil cell. *Phys. Rev. B* **37**, 6478–6481 (1988).
21. Hemley, R. J., Hanfland, M. & Mao, H. K. High pressure dielectric measurements of solid hydrogen to 170 GPa. *Nature* **350**, 488–491 (1991).
22. Garcia, H. *et al.* Dielectric properties of solid molecular hydrogen at high pressure. *Phys. Rev. B* **45**, 9809–9715 (1992).
23. Dash, W. C. & Newman, R. Intrinsic optical absorption in single crystal germanium and silicon at 77° K and 300° K. *Phys. Rev.* **99**, 1151–1155 (1955).
24. Welber, B., Cardona, M., Kim, C. K. & Rodriguez, S. Dependence of the direct energy gap of GaAs on hydrostatic pressure. *Phys. Rev. B* **12**, 5729–5738 (1975).

Acknowledgements

We thank J. Eggert for interesting discussions and suggestions.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to P.L.
(e-mail: paul.loubeyre@cea.fr).

Outgassing from Amazonian rivers and wetlands as a large tropical source of atmospheric CO₂

Jeffrey E. Richey*, John M. Melack†, Anthony K. Aufdenkampe*, Victoria M. Ballester‡ & Laura L. Hess†

* School of Oceanography, University of Washington, Seattle, Washington 98195, USA

† Institute for Computational Earth System Science, University of California, Santa Barbara, California 93106, USA

‡ Centro de Energia Nuclear na Agricultura, Caixa Postal 96, Piracicaba SP, Brazil

Terrestrial ecosystems in the humid tropics play a potentially important but presently ambiguous role in the global carbon cycle. Whereas global estimates of atmospheric CO₂ exchange indicate that the tropics are near equilibrium or are a source with respect to carbon^{1,2}, ground-based estimates indicate that the amount of carbon that is being absorbed by mature rainforests is similar to or greater than that being released by tropical deforestation^{3,4} (about 1.6 Gt C yr⁻¹). Estimates of the magnitude of carbon sequestration are uncertain, however, depending on whether they are derived from measurements of gas fluxes above forests^{5,6} or of biomass accumulation in vegetation and soils^{3,7}. It is also possible that methodological errors may overestimate rates of carbon uptake or that other loss processes have yet to be identified³. Here we demonstrate that outgassing (evasion) of CO₂ from rivers and wetlands of the central Amazon basin constitutes an important carbon loss process, equal to 1.2 ± 0.3 Mg C ha⁻¹ yr⁻¹. This carbon probably originates from organic matter transported from upland and flooded forests, which is then respired and outgassed downstream. Extrapolated across the entire basin, this flux—at 0.5 Gt C yr⁻¹—is an order of magnitude greater than fluvial export of organic carbon to the ocean⁸. From these findings, we suggest that the overall carbon budget of rainforests, summed across terrestrial and aquatic environments, appears closer to being in balance than would be inferred from studies of uplands alone^{3,5–6}.

The major biogeochemical role of river systems in the global

carbon cycle is typically considered to be the fluvial export of total organic carbon (TOC) and dissolved inorganic carbon (DIC) to the ocean (0.4–0.8 and 0.4 Gt C yr⁻¹, respectively)⁹. These fluxes are small components of the global C cycle, but they are significant compared to the net oceanic uptake of anthropogenic CO₂ (ref. 10). Aquatic carbon exports from terrestrial ecosystems, however, are not limited to fluvial discharge. Early measurements in the Amazon suggested that global CO₂ efflux (fluvial export plus respiration) from the world's rivers could be on the order of a gigatonne of carbon per year (Gt C yr⁻¹) (ref. 11). Recent measurements of mostly temperate rivers lead to estimates of global river-to-atmosphere fluxes of ~0.3 Gt C yr⁻¹, with evasion nearly equivalent to riverine TOC or DIC export¹². Evasion from streams has been estimated to average 25–50% and 28–70% of the net annual carbon accumulation for tundra¹³ and for peatlands¹⁴, respectively. Hope *et al.*¹⁴ cautioned that direct measurements of land-atmosphere CO₂ gas exchange that ignore water-borne fluxes might significantly overestimate terrestrial carbon accumulation.

The emphasis on export to the oceans and on individual streams does not, however, fully consider the spatial extent of a river network at regional scales. We evaluated the evasion of CO₂ from the fluvial environments of a 1.77 million km² quadrant of the low-gradient central Amazon basin (Fig. 1). Newly available remote sensing data sets for this quadrant have made it possible to quantify seasonal water coverage for representative low and high water periods. Combined with extensive measurements of dissolved CO₂ across the region we were able to compute the water-to-air fluxes of CO₂ for each environment.

We partitioned the quadrant into hydrographic environments—the Amazon mainstem channel, the mainstem floodplain, tributaries (channels and floodplains over 100 m in width, as constrained by the pixel dimensions of JERS-1 radar mosaics), and streams (channels and riparian zones less than 100 m in width). As computed from the radar mosaics, the flooded area of the mainstem and tributaries rose from 79,000 km² (about 4% of the quadrant area) in October 1995 to 290,000 km² (16% of the quadrant area) by May–June 1996. The low (21,000 km²) and high water (51,000 km²) areas estimated for streams were comparable to the area of the mainstem floodplain and greater than the area of the mainstem channel itself.

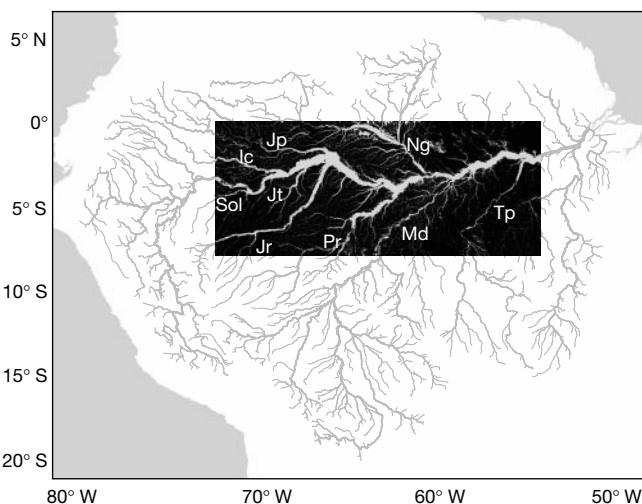


Figure 1 Flooded area of the central Amazon basin at high water, as mapped from the Japanese Earth Resources Satellite radar data (May–June 1996). The flooded area is shown as light areas in dark inset (the study quadrant). Underlying the inundation image is a digital river network (derived from the Digital Chart of the World, the GTOPO30 digital elevation model and ancillary cartographic information). Major tributaries are labelled: Negro (Ng), Japurá (Jp), Içá (Ic), Solimões (Sol, the Amazon mainstem exiting Peru), Jutai (Jt), Juruá (Jr), Purus (Pr), Madeira (Md), and Tapajós (Tp).

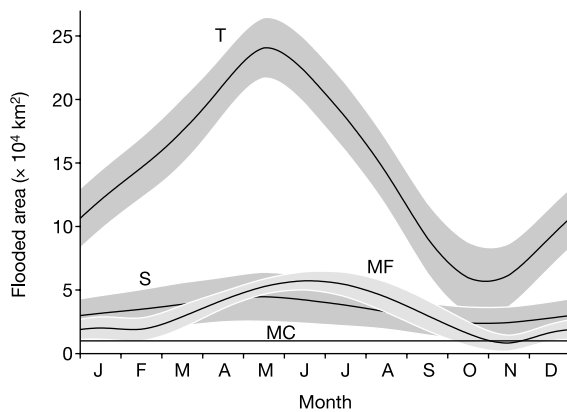


Figure 2 Spatially integrated annual sequences of surface water area for hydrographic environments as determined from JERS-1 radar data and multi-year hydrographic records. Hydrographic environments are divided into: the Amazon mainstem channel (MC); the mainstem floodplain (MF); the channels and floodplains of tributaries over 100 m wide (T); and the streams and riparian zones less than 100 m wide (S). Shaded regions represent 67% confidence intervals determined by Monte Carlo error propagation of both measurement uncertainties and interannual variability in river stage data.

Scaling to river stage height records and integrating over all sub-basins by hydrographic environment yielded the monthly extent of inundation for a typical year (Fig. 2). Tributaries south of the Amazon mainstem typically reached peak stage in April or May, whereas those to the north peaked in June or July. The result is that the region was most flooded in May (350,000 km² inundated, or 20% of the quadrant), with an annual mean flooded area of 250,000 km².

River and floodplain waters of the central Amazon basin maintain partial pressures of dissolved CO₂ (p_{CO_2}) that are supersaturated with respect to the atmosphere (Fig. 3). These concentrations track the hydrograph, increasing with rising water (to over 12,000 microatmospheres, μatm , in some tributaries) and decreasing with falling water. Average values over the year were $4,350 \pm 1,900 \mu\text{atm}$ for all mainstem samples and $5,000 \pm 3,300 \mu\text{atm}$ across the mouths of all major tributaries (compared to a mean of 3,200 μatm calculated for 47 large rivers from around the world¹²). The p_{CO_2} ranged from 2,950 μatm to over 44,000 μatm on the mainstem floodplain, averaging $14,100 \pm 10,200 \mu\text{atm}$ upstream to $6,300 \pm 4,200 \mu\text{atm}$ downstream.

High partial pressures of CO₂ translate to large gas evasion fluxes from water to atmosphere. By combining the areal extent of flooding and the distributions of p_{CO_2} with a simple but widely used gas evasion model, we observed a pronounced seasonality in evasion, corresponding to both the elevated water levels and the increased CO₂ concentrations (Fig. 4). Integrating over the year, the surface waters of the central Amazon basin (the 1.77 million km² quadrant) export $210 \pm 60 \text{ Tg C yr}^{-1}$ to the atmosphere. This corresponds to a flux of $8.3 \pm 2.4 \text{ Mg C ha}^{-1} \text{ yr}^{-1}$ over the annual mean flooded area of the central basin, or $1.2 \pm 0.3 \text{ Mg C ha}^{-1} \text{ yr}^{-1}$ over the entire quadrant.

Extrapolating across Amazonia, the total basin evasion is about 470 Tg C yr⁻¹. That is, the waters of the Amazon export about 13 times more carbon by CO₂ evasion to the atmosphere than by the export of total organic carbon (36 Tg C yr⁻¹) or of DIC (35 Tg C yr⁻¹) to the ocean⁸. Sub-basins within our study area exported 4 to 15 times more carbon by evasion than by discharge, suggesting that these findings may extend to other humid tropical river systems that drain weathered soils (and thus have low-alkalinity waters). These results are in contrast to findings from temperate river systems, where lower drainage density, higher-alkalinity waters and cooler temperatures lead to lower percentages of DIC being exported by evasion^{12,15}.

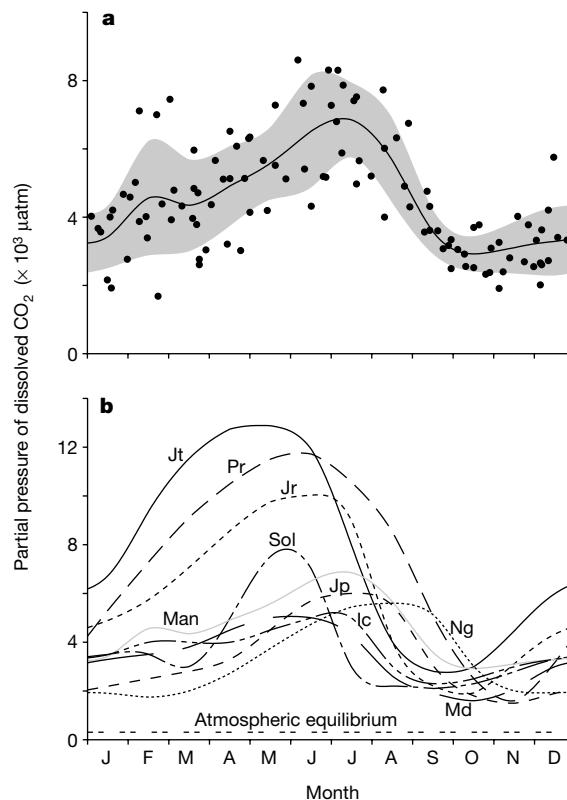


Figure 3 Seasonal distributions of carbon dioxide dissolved in rivers of the central Amazon basin. **a**, A ten-year time-series station on the mainstem Amazon near Manaus (Man), by the confluence of the Negro and Solimões offers a high-resolution image of seasonality. **b**, Annual profiles for major tributaries derived from thirteen expeditions to observation stations at the mouth of each tributary (identified in Fig. 1) describe robust patterns of spatial heterogeneity.

The large evasion raises an important ecological question: what are the carbon sources that support such a flux? We evaluated potential terrestrial and aquatic sources of CO₂, both as CO₂ and as organic carbon that could be subsequently respired¹⁶. Although the following estimates have considerable uncertainty because of large temporal and spatial variability in the attributes of the respective environments, they do indicate the relative importance of the pathways leading to evasion. The partial pressure of CO₂ in the soil atmosphere is very high (15,000 μatm at the surface to 65,000 μatm with depth at forest and pasture sites in the eastern Amazon¹⁷), resulting from root respiration and the decomposition of organic matter. Dissolution of this CO₂ and its subsequent export to streams (using a typical water through-flow rate of 1.25 m yr^{-1} ; ref. 18) is about 25% of the evasion. Dissolved organic carbon (DOC) concentrations in groundwater vary from 100 μM draining clay-rich oxisols to over 3,000 μM in groundwater draining sand-rich spodosols¹⁹. Assuming that oxisols predominate, groundwater DOC flux (computed as the product of the through-flow rate and DOC concentrations) is about 15% of the evasion. Litterfall, either directly into water or via entrainment by rising waters (estimated by applying litterfall rates for flooded forests²⁰ to the rivers and for upland forests²¹ to the streams), is about 35% of the evasion. Root respiration and decomposition from floating and emergent macrophytes, which can fix atmospheric CO₂ directly and return some of that production to water, is about 25% of the evasion (computed from annual net production²⁰ and the area of habitat from a classification of the JERS-1 images). Organic matter and CO₂ derived from the algal (phytoplankton and periphyton) and submerged macrophyte production/respiration cycles are strictly aquatic, and thus cannot be considered as potential sources.

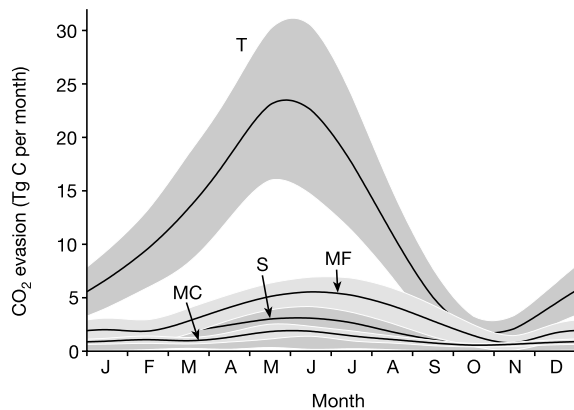


Figure 4 Spatially integrated sequences of monthly carbon dioxide evasion for the respective hydrographic environments (identified in Fig. 2). Lines represent the best estimate of long-term means, whereas shaded regions represent the 67% confidence interval for the range of values likely in a particular year. The upper confidence limit for streams, hidden from view, extends nearly to the upper limit for the mainstem floodplain.

In total, the estimate of carbon nominally available for evasion is consistent with the computed evasion. Of this amount, roughly 80% is of terrestrial upland and near-channel origin and about 20% is aquatic. Furthermore, approximately 75% originates as organic matter that would have to be respired in transit. This possibility is supported by previous work showing that DIC turnover time due to *in situ* respiration is 10–14 days for the large rivers of the Amazon²². Hence we hypothesize that evasion is driven primarily by in-stream respiration of organic carbon fixed originally on land and along river margins and mobilized into flowing waters. If true, linkages between land and water would be stronger than traditionally thought, with river corridors representing a significant downstream translocation of carbon originally fixed by the forest.

Assuming that the fluxes computed for the Amazon are representative of fluvial environments of lowland humid tropical forests in general, surface water CO₂ evasion in the tropics may help to explain an anomaly in the current balance of the global carbon cycle. Estimates that the tropics are a net carbon sink are not consistent with recent calculations from global inverse modelling, which imply that the tropics are at least in balance with the atmosphere if not a net source^{1,2}. Extrapolating over the global area covered by humid tropical forests³ with our estimate of areal evasion rates for the Amazon basin yields a flux of roughly 0.9 Gt C yr⁻¹ (three times larger than previous estimates of global evasion¹²). A return flux from water to the atmosphere of this magnitude comes closer to reconciling independent carbon budgets for the tropics. □

Methods

Areal coverage of surface waters and flooding regime

Data from the Japanese Earth Resources Satellite-1 (JERS-1) L-band synthetic aperture radar were used to estimate the areal coverage and inundation status of rivers and floodplains over 100 m in width. Data were compiled into mosaics for periods of high water (May–June 1996) and low water (October 1995)²³. For each mosaic, the study area was classified into either flooded or non-flooded areas based on radar backscatter intensities as delineated by image segmentation²⁴, and was divided into 25 tributary sub-basins from the river network (Fig. 1). The uncertainty of the procedure was estimated using high-resolution airborne, digital videography²⁵ to be $\pm 5\%$ at high water and ± 10 – 15% at low water. To account for river corridors less than 100 m in width, we computed an area density function by extending a geometric series relating stream length and width to stream order from the river network for the whole basin (Fig. 1), and applied it to the study area. Mean monthly stage data from multiyear hydrographic records within each tributary sub-basin¹⁸ were used to estimate tributary flooding sequences by assuming a temporal correspondence between stage height and areal extent of inundation. For five sub-basins without gauging records, the normalized hydrograph for the nearest neighbour with similar climatology was used. The temporal sequence of inundation within the mainstem and its floodplain was computed from multi-year monthly composite Scanning Multichannel Microwave Radiometer (SMMR) data²⁶.

CO₂ distributions

The seasonal and spatial distributions of p_{CO_2} within each hydrographic region were described from over 1,800 samples taken on 13 expeditions at different water stages throughout a 2,000 km reach of the central Amazon mainstem, tributary, and floodplain waters^{8,22,27}. A ten-year time series²² at the Marchantaria mainstem station gave a statistically robust picture of seasonal trends in p_{CO_2} (Fig. 3a). The standard deviation ranged from 16% to 37% of mean monthly values. Annual profiles for each tributary were similarly constructed, with assumed standard deviations of 40% for each month. For regions not directly sampled, profiles were used from the nearest neighbour with data, with assumed uncertainties of 80%. Floodplain measurements exhibited no obvious seasonal trends but a pronounced gradient from higher concentrations upriver to intermediate and then lower concentrations downriver. Hence we aggregated results by up-, mid-, and downriver sections.

Gas evasion

We computed evasion from the gas transfer model, $F = K(C_s - C_o)$, where F is the evasive flux, C_s is the surface water concentration, C_o is the atmospheric equilibrium, and K is the exchange coefficient. We determined K from direct measurements of O₂ and ²²²Rn accumulation in free-floating chambers on the mainstem ($K = 2.3 \pm 0.9 \text{ m d}^{-1}$) and primary tributaries ($K = 1.2 \pm 0.5 \text{ m d}^{-1}$) and of CO₂ and CH₄ accumulation on the floodplain ($K = 0.65 \pm 0.25 \text{ m d}^{-1}$)²⁸. Chambers have been criticized on the basis that they may inhibit wind shear and alter the near-surface turbulence, but we assume that in the generally low wind environments of the Amazon and with the internal turbulence of large rivers of considerable flow velocity (1–3 m s⁻¹) this model yields a reasonable, if conservative, approximation of K . These values are comparable to values derived elsewhere not only with chambers but also with purposeful injections of inert gases, calculations with surface renewal models, and eddy covariance techniques^{12,29,30}.

To extrapolate to the entire Amazon basin (to compare to total fluvial export), we needed to estimate evasion for the 4.3 million km² of the Amazon basin outside our central quadrant, for which we do not have data. This region includes not only rainforest but also drier regions of savanna and Andean headwaters, and lacks the downstream extent of the large tributaries and mainstem. It includes, however, extensive wetlands of Bolivia and the northern and eastern Amazon, which could be large sources. We assumed, probably conservatively, that the areal evasion flux outside the quadrant is half that of inside.

Uncertainty analysis

Overall 67% confidence intervals of calculated values (given as $\pm$ standard error) were determined by Monte Carlo bootstrapping error propagation of both random measurement uncertainties and interannual variability of primary variables. Variability in stage height was simulated to show considerable covariation between hydrographic regions and also within a given year for a region, and variability in p_{CO_2} was set to covary to a lesser degree. All other variability was assumed to be spatially and temporally independent. Black lines in Figs 2, 3a and 4 are thus the best estimate of long-term means, whereas shaded regions represent the range in which values for an individual month might fall 67% of the time. Systematic differences in any chosen parameter, such as K , would proportionally shift our estimates of both means and confidence limits. As in any propagation of random uncertainty, integrated or summed values exhibit lower relative uncertainty than their individual components.

Received 31 August 2001; accepted 22 February 2002.

- Schimel, D. S. *et al.* Recent patterns and mechanisms of carbon exchange by terrestrial ecosystems. *Nature* **414**, 169–172 (2001).
- Gurney, K. R. *et al.* Towards robust regional estimates of CO₂ sources and sinks using atmospheric transport models. *Nature* **415**, 626–630 (2002).
- Malhi, Y. & Grace, J. Tropical forests and atmospheric carbon dioxide. *Trends Ecol. Evol.* **15**, 332–337 (2000).
- Houghton, R. A. A new estimate of global sources and sinks of carbon from land-use change. *Eos* **81**(Suppl.), S281 (2000).
- Grace, J. *et al.* Carbon dioxide uptake by an undisturbed tropical rain forest in Southwest Amazonia, 1992 to 1993. *Science* **270**, 778–780 (1995).
- Malhi, Y. *et al.* Carbon dioxide transfer over a central Amazonian rain forest. *J. Geophys. Res.* **103**, 31593–31612 (1998).
- Phillips, O. L. *et al.* Changes in the carbon balance of tropical forests: Evidence from long-term plots. *Science* **282**, 439–442 (1998).
- Richey, J. E. *et al.* Biogeochemistry of carbon in the Amazon River. *Limnol. Oceanogr.* **35**, 352–371 (1990).
- Degens, E. T., Kempe, S. & Richey, J. E. (eds) *Biogeochemistry of Major World Rivers* 323–397 (Wiley, Chichester, 1991).
- Sarmiento, J. L. & Sundquist, E. T. Revised budget for the oceanic uptake of anthropogenic carbon dioxide. *Nature* **356**, 589–593 (1992).
- Richey, J. E., Brock, J. T., Naiman, R. J., Wissmar, R. C. & Stallard, R. F. Organic carbon: oxidation and transport in the Amazon River. *Science* **207**, 1348–1351 (1980).
- Cole, J. J. & Caraco, N. F. Carbon in catchments: Connecting terrestrial carbon losses with aquatic metabolism. *Mar. Freshwat. Res.* **52**, 101–110 (2001).
- Kling, G. W., Kipphut, G. W. & Miller, M. C. Arctic lakes and streams and gas conduits to the atmosphere: implications for tundra carbon budgets. *Science* **251**, 298–301 (1991).
- Hope, D., Palmer, S. M., Billett, M. F. & Dawson, J. J. Carbon dioxide and methane evasion from a temperate peatland stream. *Limnol. Oceanogr.* **46**, 847–857 (2001).
- Telmer, K. & Veizer, J. Carbon fluxes, p_{CO_2} and substrate weathering in a large northern river basin, Canada: carbon isotope perspectives. *Chem. Geol.* **159**, 61–86 (1999).
- McClain, M. E. & Richey, J. E. Regional-scale linkages of terrestrial and lotic ecosystems in the Amazon basin: A conceptual model for organic matter. *Arch. Hydrobiol.* **113**, (Suppl.), 111–125 (1996).

17. Davidson, E. A. & Trumbore, S. E. Gas diffusivity and production of CO₂ in deep soils of the eastern Amazon. *Tellus B* **47**, 550–565 (1995).
18. Richey, J. E., Victoria, R. L., Mayorga, E., Martinelli, L. A. & Meade, R. H. in *Biospheric Feedbacks in Climate and the Hydrological Cycle* (ed. Kabat, P.) (Springer, in the press).
19. McClain, M. E., Richey, J. E., Brandes, J. A. & Pimentel, T. P. Dissolved organic matter and terrestrial-lotic linkages in the central Amazon Basin, Brazil. *Glob. Biogeochem. Cycles* **11**, 295–311 (1997).
20. Melack, J. M. & Forsberg, B. R. in *The Biogeochemistry of the Amazon Basin* (eds McClain, M. E., Victoria, R. L. & Richey, J. E.) 235–274 (Oxford Univ. Press, New York, 2001).
21. Chambers, J. Q., dos Santos, J., Ribeiro, R. J. & Higuichi, N. Tree damage, allometric relationships, and above-ground net primary production in central Amazon forest. *Forest Ecol. Management* **5348**, 1–12 (2000).
22. Devol, A. H., Forsberg, B. R., Richey, J. E. & Pimentel, T. P. Seasonal variation in chemical distributions in the Amazon (Solimões) River: a multiyear time series. *Glob. Biogeochem. Cycles* **9**, 307–328 (1995).
23. Siqueira, P. et al. A continental-scale mosaic of the Amazon Basin using JERS-1 SAR. *IEEE Trans. Geosci. Remote Sensing* **38**, 2638–2644 (2000).
24. Barbosa, C., Hess, L., Melack, J. & Novo, E. Mapping Amazon Basin wetlands through region-growing segmentation and segmented-based classification of JERS-1 data. *IX Latin Am. Symp. Remote Sensing (6–10 November 2000)* 1168–1176 (Universidad Nacional de Lujan, Puerto Iguazu, Argentina, 2000); see also (<http://www.selper.org>).
25. Hess, L. L. et al. Geocoded digital videography for validation of land cover mapping in the Amazon Basin. *Int. J. Remote Sensing* (in the press).
26. Sippel, S. J., Hamilton, S. K., Melack, J. M. & Novo, E. M. Passive microwave observations of inundation area and the area/stage relation in the Amazon River floodplain. *Int. J. Remote Sensing* **19**, 3055–3074 (1998).
27. Richey, J. E., Devol, A. H., Wofsy, S. C., Victoria, R. & Ribeiro, M. N. G. Biogenic gases and the oxidation and reduction of carbon in the Amazon River and floodplain waters. *Limnol. Oceanogr.* **33**, 551–561 (1988).
28. Devol, A. H., Quay, P. D., Richey, J. E. & Martinelli, L. A. The role of gas exchange in the inorganic carbon, oxygen and 222 radon budgets of the Amazon River. *Limnol. Oceanogr.* **32**, 235–248 (1987).
29. Clark, J. F., Wanninkhof, R., Schlosser, P. & Simpson, H. J. Gas exchange rates in the tidal Hudson River using a dual tracer technique. *Tellus B* **46**, 264–285 (1994).
30. MacIntyre, S., Eugster, W. & Kling, G. W. in *Gas Transfer at Water Surfaces* (eds Donelan, M. A., Drennan, W. M., Saltzman, E. S. & Wanninkhof, R.) 135–139 (American Geophysical Union, Washington, 2001).

Acknowledgements

We thank E. Mayorga, S. Denning, M. Gastil, D. Montgomery, R. Victoria, A. Krusche, A. Devol, P. Quay and J. Hedges for technical assistance and discussions, B. Forsberg and T. Pimentel for fieldwork, and the Global Rain Forest Mapping Project of the National Space Development Agency of Japan for providing the JERS-1 radar data. This work was supported by the US NSF and NASA EOS and LBA projects, and by the Brazilian FAPESP programme.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to J.E.R. (e-mail: jrichey@u.washington.edu).

Small-scale structure of the geodynamo inferred from Oersted and Magsat satellite data

Gauthier Hulot*, Céline Eymin*, Benoît Langlais*, Mioara Mandea* & Nils Olsen†

* Département de Géomagnétisme et Paléomagnétisme, CNRS UMR 7577, Institut de Physique du Globe de Paris, 4 Place Jussieu, B89, Tour 24, 75252 Paris cedex 05, France

† Center for Planetary Science, Danish Space Research Institute, Juliane Maries Vej 30, DK-2100 Copenhagen, Denmark

The 'geodynamo' in the Earth's liquid outer core produces a magnetic field that dominates the large and medium length scales of the magnetic field observed at the Earth's surface^{1,2}. Here we use data from the currently operating Danish Oersted³ satellite, and from the US Magsat² satellite that operated in 1979/80, to identify and interpret variations in the magnetic field over the past 20 years, down to length scales previously inaccessible.

Projected down to the surface of the Earth's core, we found these variations to be small below the Pacific Ocean, and large at polar latitudes and in a region centred below southern Africa. The flow pattern at the surface of the core that we calculate to account for these changes is characterized by a westward flow concentrated in retrograde polar vortices and an asymmetric ring where prograde vortices are correlated with highs (and retrograde vortices with lows) in the historical (400-year average) magnetic field^{4,5}. This pattern is analogous to those seen in a large class of numerical dynamo simulations⁶, except for its longitudinal asymmetry. If this asymmetric state was reached often in the past, it might account for several persistent patterns observed in the palaeomagnetic field^{7–10}. We postulate that it might also be a state in which the geodynamo operates before reversing.

Thanks to the recent launch of the Danish Oersted satellite³ (inclination 96.5°, altitude 638–849 km), 20 years after the 1979/80 US Magsat² analogous mission (97°, 325–550 km), two data sets at two different epochs are now available that can be used to construct high-degree spherical harmonic models of the geomagnetic field. (Degree 1 is the dipole field; the larger the degree, the smaller the length scale.) Taking advantage of this opportunity and relying on models¹¹ well suited for that purpose, we compute and investigate the changes that have occurred in the geomagnetic field between 1980 and 2000, focusing on the large to medium scales (that is, up to

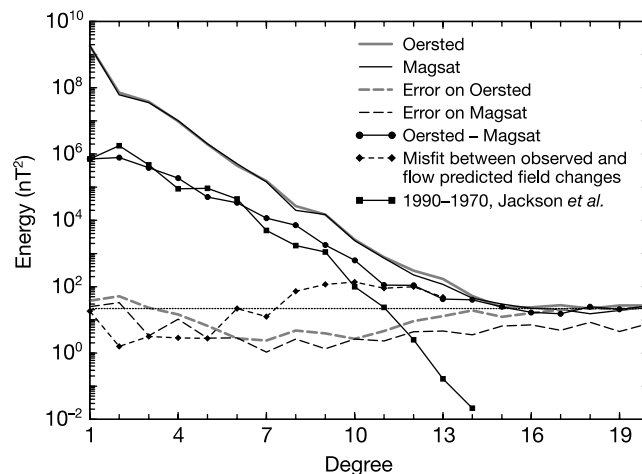


Figure 1 Spectra of the Oersted and Magsat models¹¹, of their error, of their difference, and of how well flows predict this difference. The spectrum of a model is constructed by plotting the contribution of each degree n of the spherical harmonic expansion to the average $\langle B^2 \rangle$ of the predicted field B over the Earth's surface^{1,2}. Errors in the Oersted and Magsat models are computed by analysing differences between models based on various data subsets¹¹. Note the well-known² knee within the Oersted and Magsat spectra, showing that the main field probably dominates the signal for degrees less than 13, whereas the crustal field dominates for degrees larger than 15. A similar knee is seen around degree 15, in the Oersted–Magsat spectrum of the difference between the Oersted and Magsat fields. This knee, and the flat section of the spectrum beyond it, reveals disagreements between the high-degree signal sensed by Oersted and Magsat at a level slightly less than that of the crustal signal itself¹². Up to degree 13, however, the Oersted–Magsat spectrum is well above the level (dotted line) defined by both that flat section and the crustal spectrum (the contribution of which is believed to be weaker at low degrees than at high degrees³). It is also well above the error level. It thus cannot be attributed to noise or a crustal source. By contrast, it can be explained by core surface flows, as is illustrated by the spectrum of the misfit between the predicted and the observed Oersted minus Magsat field difference. This misfit is at a level comparable to that of the errors in the models and of crustal contributions, but relaxed at the largest degrees to account for 'truncation errors'¹⁷. For reference, the spectrum of the less-resolved field variations⁵ between 1970 and 1990 is also shown.

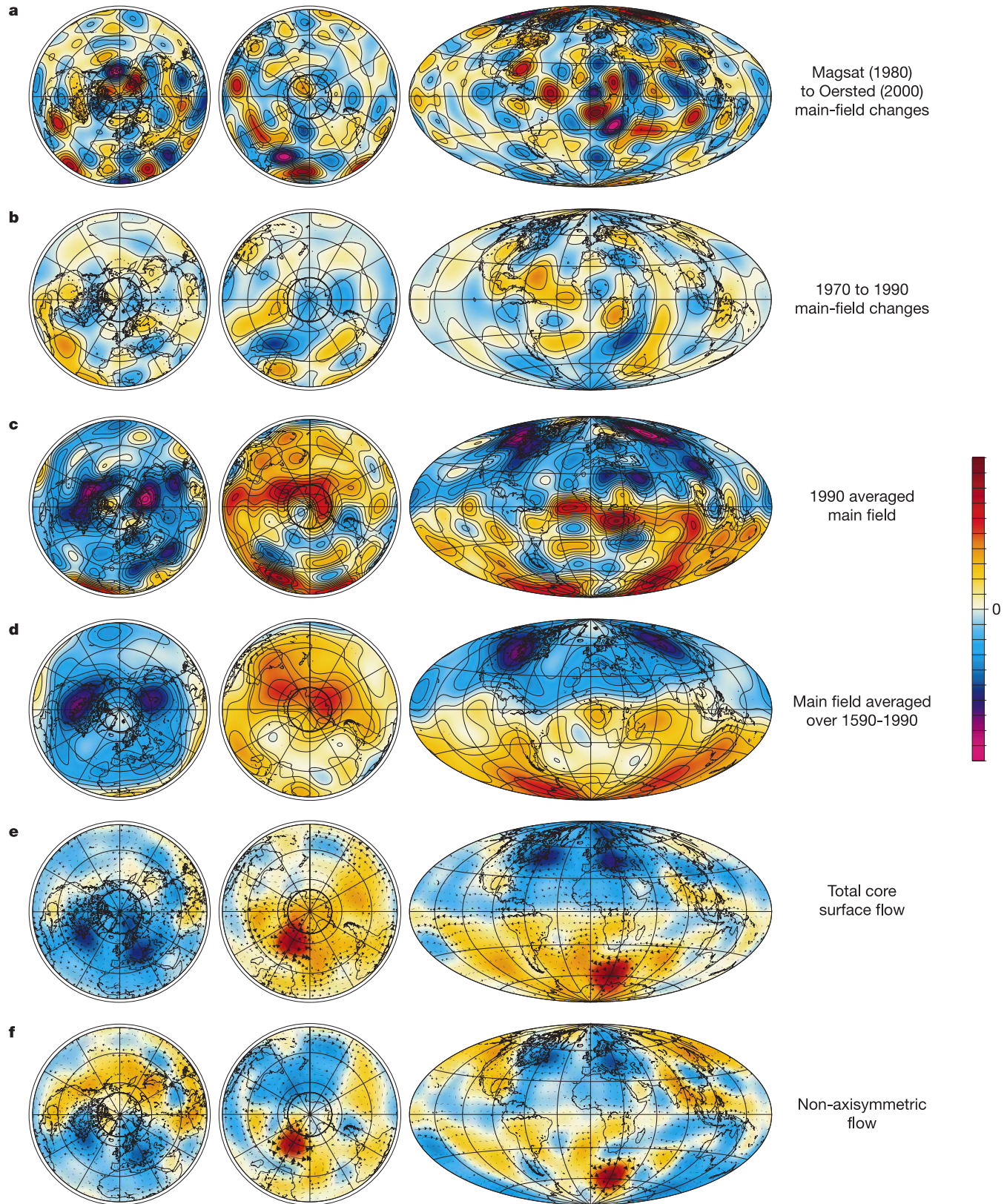


Figure 2 Polar (north and south) and Hammer views of the small-scale structure of the geodynamo at the core surface. **a**, Changes in the radial component of the field between 1980 (Magsat) and 2000 (Oersted) ($-340 \mu\text{T}$ to $333 \mu\text{T}$). **b**, Analogous but less-resolved field changes between 1970 and 1990 (computed from the historical model of ref. 5, $-186 \mu\text{T}$ to $147 \mu\text{T}$). **c**, The radial component of the average 1990 main field computed by averaging the Magsat and Oersted main-field models¹¹ ($-1,032 \mu\text{T}$ to $908 \mu\text{T}$). **d**, The main field averaged over the historical period 1590–1990 (computed from the model of ref. 5, $-745 \mu\text{T}$ to $594 \mu\text{T}$). **e**, Core surface flow accounting for the

main-field changes in **a** by advecting the 1990 main field in **c** (arrows for the flow, maximum of 50 km yr^{-1} , colour code for the toroidal scalar associated to the flow). **f**, Same as in **e** but for the non-axisymmetric component of the flow (maximum 65 km yr^{-1}). In all figures, there is a linear colour code: red positive, blue negative, renormalized to the maximum absolute value, except for **b** (respectively **d**) which uses the same scale as **a** (respectively **c**). Contours every $50 \mu\text{T}$ in **a** and **b**, every $100 \mu\text{T}$ in **c** and **d**. Also shown in each polar plot, the surface trace of the tangent cylinder.

degree 13) thought to be dominated by the main field produced by the geodynamo^{1,2} (Fig. 1).

Figure 2a shows the radial component of those changes plotted at the core surface where the main field originates. Most remarkable are the large changes, which occurred at high, especially northern, latitudes, and in a hemispheric region centred below Africa. These regions contrast strongly with a wide region below the Pacific, where far fewer changes occurred. This pattern could not be resolved as clearly before the Oersted mission (ref. 5, Fig. 2b), and is robust against model uncertainties (Fig. 1).

We now consider if this pattern could have been produced by sources other than the geodynamo—perhaps in the crust, ionosphere or magnetosphere. Although formal separation of the signals from these sources is only possible in part, over many years a framework for practical separation has been constructed², and the results of such separation suggest that this could hardly be the case. First, we consider crustal signals. These are thought to dominate the field at degree 15 and above (Fig. 1). But there is no evidence at these wavelengths for any large changes in the field between 1980 and 2000¹², and at longer wavelengths, even a 100% change in the estimated crustal signal would produce weak field changes (Fig. 1). Second, ionospheric signals. In the field models on which we rely¹¹, these signals have been carefully minimized by using night-side measurements on magnetically quiet days, and field-aligned currents at high latitudes have specifically been removed by using only field intensity data there. The remaining ionospheric contamination at the Earth's surface can then be estimated to be less than 10 nT at all latitudes in the Magsat (1980) model², and probably less in the Oersted (2000) model. Much of this field would in addition show up in zonal fields in Fig. 2a, as the data for each model were recorded at similar local times. Such characteristics do not match the typical 100-nT change (mainly in the degrees 8 to 13 and dominantly not in zonal fields) implied at the Earth's surface by the intense small-scale structures seen in Fig. 2a. Magnetospheric signals can also easily be dismissed, as they are large-scale and can be formally separated from the core signal². Finally, signals from electrical currents induced in the crust and upper mantle by all those external sources can again be shown to be too weak^{13,14}. Therefore, the results in Fig. 2a are most likely to be due to the temporal (secular) variation of the main field.

This secular variation is the combined consequence of main-field diffusion through ohmic dissipation and main-field advection by flows at the core surface¹⁵. Previous studies^{16,17} have shown that much of the short-term large-scale variations of the main field could be explained in terms of its advection by surface flows satisfying the tangentially geostrophic balance¹⁸ (which assumes that the horizontal component of the Coriolis forces is mainly balanced by a dynamical pressure gradient), in the so-called 'frozen-flux approximation'¹⁵ (which assumes diffusion to be negligible). Core flows inferred in this way have further been shown to account for length of day variations on decade timescales (by angular momentum exchange with the mantle)^{19–21}, and to reflect dynamical features, akin to torsional oscillations predicted by dynamo theory^{22,21}. We thus decided to try and compute a tangentially geostrophic flow, which would account for the observed main-field changes (Fig. 2a).

To do this, we computed the average of the two 1980 and 2000 main-field models (Fig. 2c), together with the average secular variation accounting for the 1980 to 2000 main-field changes. We finally computed the flow producing this average secular variation by advection of the average main-field model. We relied on the numerical procedure of ref. 21 with some adjustments. More details about this computation will be provided elsewhere (C.E. and G.H., manuscript in preparation). Most important here is the fact that the flow that we computed (Fig. 2e) succeeds at predicting the field changes at a satisfactory level (see Fig. 1).

Figure 2e reveals a flow with a number of strong vortices embedded in a mainly westward axisymmetric flow. This axisymmetric flow (detailed in Fig. 3) is remarkably symmetric with respect

to the equator. It consists mainly of a westward body rotation of the core with respect to the mantle (of order 0.1°yr^{-1}) and two strong westward ('retrograde' when compared to the Earth's daily rotation) polar vortices (of order 0.9°yr^{-1} , the northern vortex being slightly larger) within the 'tangent cylinder' (tangent to the inner core and intersecting the core surface at latitude $\pm 69.5^\circ$). The body rotation is a well-known feature. It varies with time as a result of core–mantle coupling acting on decade timescales^{16,19–21}. Far less is known about the polar vortices, which confine most of the westward drift at high latitude and appear to be important in explaining the changes seen near both poles in Fig. 2a. Recently, similar vortices have been tentatively found with the help of computations based on less-resolved historical models^{21,23}. These vortices may thus have been active over at least the past century, and could be permanent features of the geodynamo. Figure 3 also shows a small-scale zonal flow that we believe is the result of a Gibbs effect produced by the flow computation, which still fails to resolve the finest details of the probably sharp boundaries of the polar vortices²¹. We finally note that no significant medium-scale zonal flow is present in Fig. 3. This suggests that the torsional oscillations are currently of low amplitude, a result consistent with both the general trend seen in core flows computed over the past century²¹ and the relatively short decay time which has been inferred for those oscillations²².

Plotting the non-axisymmetric flow alone leads to another interesting result (Fig. 2f). This part of the flow displays a medium-to-high-latitude ring of vortices roughly, but not exactly, symmetrical with respect to the equator, and clustering around the tangent cylinder. This is best seen by also plotting the toroidal scalar associated with the flow, which can be viewed as a filtered (low-pass) measure of the radial vorticity $\omega_r = \mathbf{n} \cdot (\nabla \times \mathbf{u})$ of the flow $\mathbf{u}$. The colour code in Fig. 2e and f is then such that prograde and retrograde vortices appear respectively red and blue in the Northern Hemisphere, and respectively blue and red in the Southern Hemisphere (note that equatorial-symmetric flows translate into anti-symmetric radial vorticity). Computations based on historical models could not resolve these vortices as clearly^{16,17}. They are reminiscent of analogous vortices seen in many numerical simulations of the geodynamo^{6,24,25}. The class of simulations that are referred to in ref. 6 as being in a 'fully developed regime' (see also ref. 24) also display retrograde polar vortices analogous to those

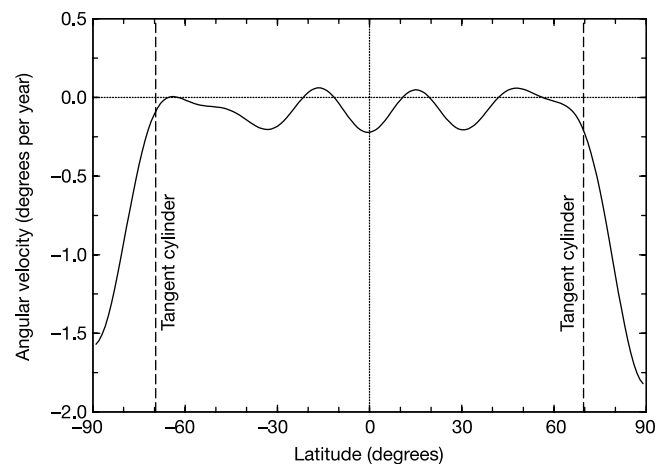


Figure 3 Axisymmetric component of the flow at the core surface. This azimuthal flow is plotted in terms of angular velocity about the Earth's rotation axis (positive for eastward with respect to the mantle, negative for westward) as a function of latitude. Note the general westward trend, and the strong westward flows at high latitudes. These flows correspond to retrograde polar vortices occurring within the tangent cylinder (beyond $\pm 69.5^\circ$ of latitude); see text.

identified in Fig. 3. Also, for such dynamos, computing surface flows as we did has been shown to lead to a reasonable surface picture of the full dynamo flows, provided those flows are not too small scale²⁶. Even though it is known that such numerical simulations are still forced to use unrealistic dimensionless numbers²⁷, this suggests that the geodynamo could belong to a similar 'class' of dynamos.

For such dynamos, it has also been observed that prograde vortices are associated with slow downwelling flows, which concentrate the flux inside the vortices on the long term. By contrast, retrograde vortices are associated with slow upwelling flows, which tend to expel field of reversed polarity (compared to that of the main dipole field) from within the core. It has been speculated that a similar phenomenon could take place in the Earth's core²⁸. In Fig. 2f, prograde vortices (red in the North, blue in the South) tend to occur in a region where maxima are seen in the main field (where the so-called 'flux bundles'^{4,28} are found, Fig. 2c). By contrast, retrograde vortices (blue in the North, red in the South) occur where the main field tends to be minimum, if not of reversed polarity. This correlation is much clearer when made with the main field averaged over the past 400 years (Fig. 2d). Analogous up- and downwelling flows (weak and non-geostrophic, hence not directly visible in Fig. 2f which shows the first-order flow computed under the tangentially geostrophic assumption) could thus be responsible for the global structure of the main field averaged over secular timescales.

One characteristic of the geodynamo remains puzzling: its azimuthal asymmetry. In most numerical dynamos, prograde and retrograde vortices tend to alternate around the tangent cylinder. In Fig. 2f, the flow in the 'Pacific hemisphere' is mainly made up of prograde vortices, contrasting with the flow in the other hemisphere where strong retrograde vortices are found. This is also where the strongest field changes are seen at present (Fig. 2a), and where most of the reversed polarity field has been produced in the past 400 years (which led to the creation of the large reverse patch now seen below South Africa in Fig. 2c)⁴. Archaeomagnetic data further suggest that significant changes have occurred in the (large-scale) main field over the past 3,000 years (ref. 29), and that the present main-field pattern is a relatively recent feature. This leads us to speculate that in the past millennium, retrograde vortices could have progressively vanished in the Pacific hemisphere, while gaining momentum and increasing the rate of creation of reversed polarity field in the other hemisphere. At the present time, the changes that this would have produced in the main field would still be affected by the strong main flow associated with the vortices, leading to the locally enhanced field changes seen in Fig. 2a. It could thus be that the asymmetry at present observed in all maps of Fig. 2 is only temporary. This would be consistent with the results of the only dynamo explicitly displaying a similar azimuthal asymmetry, which is indeed only temporarily observed²⁵.

Palaeomagnetic data can then be used to try and gain further insight into this question. The present asymmetric main field is also associated with a significantly enhanced order 1 (that is, $\cos \phi$) non dipole field component^{1,7}. If this were to happen often enough, it could explain most of the behaviour of the palaeosecular variation over the past 5 Myr (refs 1, 7). This then further suggests that, although possibly temporary, the present asymmetric state of the geodynamo could be frequently reached. If this state were preferentially reached in a fixed way with respect to the mantle, as a result of its likely influence on the core, it could also eventually produce an averaged palaeomagnetic field showing traces of 'flux bundles'^{4,28} as in Fig. 2d. But whether such traces can in fact be seen in the palaeomagnetic field remains an open question^{8,30}.

The growth of the South African patch could also be associated with the present rapid decrease of the dipole field—a decrease that some^{9,10} have suggested could eventually lead to a reversal sharing characteristics with the last known reversal (780 kyr ago¹). Although this does not prove that a new reversal is impending, it suggests

that the asymmetric state we are witnessing at present is one through which the geodynamo could also possibly go just before reversing. □

Received 4 December 2001; accepted 4 February 2002.

- Merrill, R. T., McElhinny, M. W. & McFadden, P. L. *The Magnetic Field of the Earth, Paleomagnetism, the Core and the Deep Mantle* (Academic, San Diego, 1996).
- Langel, R. A. & Hinze, W. J. *The Magnetic Field of the Earth's Lithosphere, The Satellite Perspective* (Cambridge Univ. Press, Cambridge, 1998).
- Neubert, T. et al. Oersted satellite captures high-precision geomagnetic field data. *Eos* **82**, 81, 87–88 (2001).
- Bloxham, J., Gubbins, D. & Jackson, A. Geomagnetic secular variation. *Phil. Trans. R. Soc. Lond. A* **329**, 415–502 (1989).
- Jackson, A., Jonkers, A. R. T. & Walker, M. R. Four centuries of geomagnetic secular variation from historical records. *Phil. Trans. R. Soc. Lond. A* **358**, 957–990 (2000).
- Olson, P., Christensen, U. & Glatzmaier, G. Numerical modeling of the geodynamo: Mechanisms of field generation and equilibration. *J. Geophys. Res.* **104**, 10383–10404 (1999).
- Hulot, G. & Gallet, Y. On the interpretation of virtual geomagnetic pole (VGP) scatter curves. *Phys. Earth. Planet. Inter.* **95**, 37–53 (1996).
- Gubbins, D. & Kelly, P. Persistent pattern in the geomagnetic field over the past 2.5 Myr. *Nature* **365**, 829–832 (1993).
- Gubbins, D. & Coe, R. Longitudinally confined geomagnetic reversal paths from non-dipole transition fields. *Nature* **362**, 51–53 (1993).
- Constable, C. Link between geomagnetic reversal paths and secular variation of the field over the past 5 Myr. *Nature* **358**, 230–233 (1992).
- Langlais, B., Manda, M. & Ultré-Guérard, P. High-resolution magnetic field modelling: application to MAGSAT and Oersted data. *Phys. Earth. Planet. Inter.* (in the press).
- Purucker, M., Langlais, B., Olsen, N., Hulot, G. & Manda, M. The southern edge of cratonic North America: Evidence from new satellite magnetometer observations. *Geophys. Res. Lett.* (in the press).
- Olsen, N. Induction studies with satellite data. *Surv. Geophys.* **20**, 309–340 (1999).
- Tarits, P. & Naphsica, G. Electromagnetic induction effects by the solar quiet magnetic field at satellite altitude. *Geophys. Res. Lett.* **27**, 4009–4012 (2000).
- Roberts, P. H. & Scott, S. On the analysis of the secular variation. 1. A hydromagnetic constraint: Theory. *J. Geomagn. Geoelectr.* **17**, 137–151 (1965).
- Bloxham, J. & Jackson, A. Fluid flow near the surface of Earth's outer core. *Rev. Geophys.* **29**, 97–120 (1991).
- Hulot, G., Le Mouél, J. L. & Wahr, J. Taking into account truncation problems and geomagnetic model accuracy in assessing computed flows at the core-mantle boundary. *Geophys. J. Int.* **108**, 224–246 (1992).
- Le Mouél, J. L. Outer core geostrophic flow and secular variation of Earth's magnetic field. *Nature* **311**, 734–735 (1984).
- Jault, D., Gire, C. & Le Mouél, J. L. Westward drift, core motions and exchanges of angular momentum between core and mantle. *Nature* **333**, 353–356 (1988).
- Jackson, A., Bloxham, J. & Gubbins, D. in *Dynamics of Earth's Deep Interior and Earth Rotation* (eds Le Mouél, J.-L. et al.) 97–107 (IUGG Vol. 12, AGU Geophysical Monograph 72, American Geophysical Union, Washington DC, 1993).
- Pais, A. & Hulot, G. Length of day decade variations, torsional oscillations and inner core superrotation: evidence from recovered core surface zonal flows. *Phys. Earth. Planet. Inter.* **118**, 291–316 (2000).
- Zatman, S. & Bloxham, J. Torsional oscillations and the magnetic field within the Earth's core. *Nature* **388**, 760–763 (1997).
- Olson, P. & Aurnou, J. A polar vortex in the Earth's core. *Nature* **402**, 170–173 (1999).
- Glatzmaier, G. & Roberts, P. A three-dimensional convective dynamo solution with rotating and finitely conducting inner core and mantle. *Phys. Earth. Planet. Inter.* **91**, 63–75 (1995).
- Kuang, W. & Bloxham, J. in *The Core-Mantle Boundary Region* (eds Gurnis, M. et al.) 187–208 (Geodyn. Ser. 28, American Geophysical Union, Washington DC, 1998).
- Rau, S., Christensen, U., Jackson, A. & Wicht, J. Core flow inversion tested with numerical dynamo models. *Geophys. J. Int.* **141**, 485–497 (2000).
- Dormy, E., Valet, J. P. & Courtillot, V. Numerical models of the geodynamo and observational constraints. *Geochem. Geophys. Geosyst.* **1**, 62 (2000).
- Gubbins, D. & Bloxham, J. Morphology of the geomagnetic field and implications for the geodynamo. *Nature* **325**, 509–511 (1987).
- Constable, C. G., Johnson, C. L. & Lund, S. P. Global geomagnetic field models for the past 3000 years: transient or permanent flux lobes? *Phil. Trans. R. Soc. Lond. A* **358**, 991–1008 (2000).
- Carlut, J. & Courtillot, V. How complex is the Earth's average magnetic field? *Geophys. J. Int.* **134**, 527–544 (1998).

Acknowledgements

We thank those involved in the Oersted project for their contribution to this work. The Oersted project is funded by the Danish Ministry of Transport, the Ministry of Research and Information Technology, and the Ministry of Trade and Industry of Denmark. Additional support for Oersted came from NASA, the Centre National d'Etudes Spatiales (CNES) and DARA. We also thank R. Holme for comments and suggestions.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to G.H. (e-mail: gh@ipgp.jussieu.fr).

Strong male-driven evolution of DNA sequences in humans and apes

Kateryna D. Makova & Wen-Hsiung Li

Department of Ecology and Evolution, University of Chicago, 1101 East 57th Street, Chicago, Illinois 60637, USA

Studies of human genetic diseases have suggested a higher mutation rate in males than in females¹ and the male-to-female ratio (α) of mutation rate has been estimated from DNA sequence and microsatellite data to be about 4–6 in higher primates^{2–5}. Two recent studies, however, claim that α is only about 2 in humans^{6,7}. This is even smaller than the estimates ($\alpha > 4$) for carnivores and birds^{8,9}; humans should have a higher α than carnivores and birds because of a longer generation time and a larger sex difference in the number of germ cell cycles. To resolve this issue, we sequenced a noncoding fragment on Y of about 10.4 kilobases (kb) and a homologous region on chromosome 3 in humans, greater apes, and lesser apes. Here we show that our estimate of α from the internal branches of the phylogeny is 5.25 (95% confidence interval (CI) 2.44 to ∞), similar to the previous estimates^{2–5}, but significantly higher than the two recent ones^{6,7}. In contrast, for the external (short, species-specific) branches, α is only 2.23 (95% CI: 1.47–3.84). We suggest that closely related species are not suitable for estimating α , because of ancient polymorphism and other factors. Moreover, we provide an explanation for the small estimate of α in a previous study⁶. Our study reinstates a high α in hominoids and supports the view that DNA replication errors are the primary source of germline mutation.

An effective way to estimate α is to use highly similar noncoding sequences on different types of chromosomes¹⁰. Noting that the *DAZ* locus was translocated from chromosome 3 to Y after the split between Old and New World monkeys¹¹, we sequenced three noncoding segments (~10.4 kb total) on chromosomes Y and 3 in the human, bonobo, gorilla, siamang and gibbon.

When we estimate the α -values separately for the external and the

internal branches of the tree (Fig. 1) an intriguing pattern emerges (Table 1). The α -values vary enormously among the external branches, but their average is low. The sum of the external branch lengths is 4.43% for the Y sequences (Y) and 3.22% for the chromosome 3 sequences (A), leading to $Y/A = 1.38$. According to a previous study¹²: $Y/A = 2\alpha/(1 + \alpha)$, so $\alpha = 2.23$ (95% CI: 1.47–3.84), not significantly different from the estimate (1.55) in ref. 6. In contrast, the sum of the internal branch lengths is 3.77% for the Y sequences and 2.25% for the chromosome 3 sequences, leading to $Y/A = 1.68$ and $\alpha = 5.25$ (95% CI: 2.44 to ∞). This is similar to the earlier estimates of α in primates^{2–5}, but significantly higher than those in refs 6 and 7. Thus, estimates of α obtained from closely related species tend to be lower than those obtained from more distantly related species.

We wondered how to explain the above discrepancy. The level of polymorphism on chromosome Y is usually extremely low because of selective sweep and background selection^{13,14} and because the effective population size of Y is only one-quarter of that of an autosome. The average divergence between two species is equal to $\pi + 2ut$, where t is the divergence time, u is the mutation rate, and π is the average divergence between two sequences in the ancestral population (ancient nucleotide diversity) at the time of speciation¹⁵. The ratio of Y divergence (Y) to autosome divergence (A) is

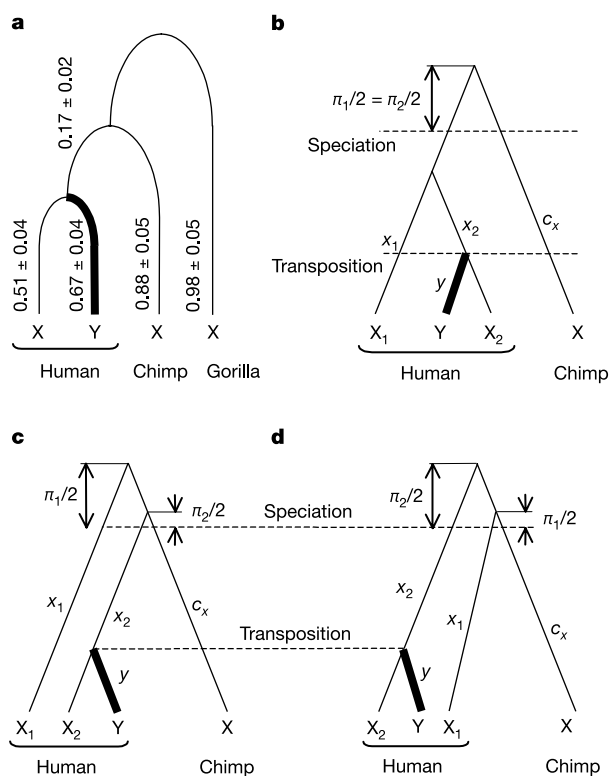


Figure 2 Phylogenetic schemes for the sequences of ref. 6. **a**, The original scheme (modified from ref. 6). **b–d**, Three possible schemes: one of them is likely to be true. π_1 , the ancient nucleotide diversity between the ancestral human X_1 and chimpanzee X chromosomes; π_2 , the ancient nucleotide diversity between the ancestral human X_2 and chimpanzee X chromosomes; x_1 , the number of nucleotide substitutions on human X_1 since the common ancestor of X_1 and X_2 or since the speciation; x_2 , the number of nucleotide substitutions on human X_2 between the common ancestor of X_1 and X_2 (or the speciation) and the transposition; c_x , the number of nucleotide substitutions on chimpanzee X chromosome; y , the number of nucleotide substitutions on the branch for the human Y chromosome locus from the time of transposition to present. Note that **a** assumes $X_2 = X_1$.

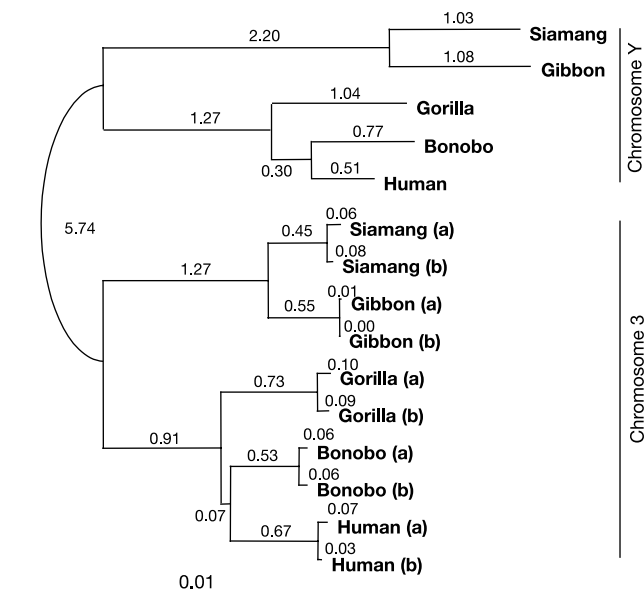


Figure 1 Phylogenetic tree of nucleotide sequences. Branch lengths were estimated from the pairwise evolutionary distances (substitutions per 100 sites). The pseudosequences of the two alleles at the chromosome 3 locus in each species are labelled a and b.

Table 1 Male-to-female ratio (α) of mutation rate

Branch	Y*	A*	Y/A	α	95% CI
External branches					
Human	0.51	0.72	0.71	0.55	0.31–0.98
Bonobo	0.77	0.59	1.31	1.90	0.87–13.29
Gorilla	1.04	0.83	1.25	1.67	0.90–5.50
Siamang	1.03	0.52	1.98	99.00	2.40– ∞
Gibbon	1.08	0.56	1.93	27.57	2.40– ∞
Total	4.43	3.22	1.38	2.23	1.47–3.84
Internal branches	3.77	2.25	1.68	5.25	2.44– ∞

Ratios were estimated for external (species-specific) and internal branches from a ~10.4-kb segment.

* Branch length in terms of the number of substitutions per 100 sites.

$Y/A = (y + \pi_y)/(a + \pi_a)$, where y is the number of mutations on Y, a is the number of mutations on an autosome, and π_y and π_a are the ancient nucleotide diversities at the point of speciation on Y and an autosome, respectively. As π_y is usually lower than π_a ($y + \pi_y)/(a + \pi_a)$ can be substantially lower than y/a for closely related species. For instance, π_a between the two gorilla alleles at the chromosome 3 locus is 0.19% (Fig. 1). If we assume $\pi_a = 0.19\%$ and $\pi_y \approx 0$ in the common ancestor of the bonobo and the gorilla, then Y/A between the bonobo and the gorilla increases from $2.11\%/1.49\% = 1.42$ to $2.11\%/(1.49\% - 0.19\%) = 1.62$, and α increases from 2.45 to 4.26. Clearly, ancient polymorphism can considerably reduce the estimate of α for closely related species. If A is small, even a small error in the estimate of A can have a strong effect on Y/A .

A more serious problem in ref. 6 is that their phylogeny (Fig. 2a) assumes that the Y-linked sequence, which was derived from the transposition of an X-linked sequence in an ancestral human population, has always evolved as a Y-linked sequence since its separation from the X-linked sequence sampled in their study. This is unlikely. Rather, one of the schemes in Fig. 2b–d should be true. In Fig. 2b, when chimpanzee X is used as a reference, we have $X = x_1$ and $Y = x_2 + y$. If $y < x_2$, Y can be close to X and α can be seriously underestimated. In Fig. 2c, $X = \pi_1 + x_1$, $Y = \pi_2 + x_2 + y$ and $\pi_1 > \pi_2$, so if y is very small, Y can even be smaller than X ; that is, both Y/X and α can be smaller than 1. In Fig. 2d, $\pi_1 < \pi_2$, so Y/X is greater than 1. However, Y/X can be close to 1, if y is small and if π_2 is close to π_1 . Among the three schemes, Fig. 2b is most probable because the observed human X–human Y divergence (1.18%) is smaller than the observed human X–chimpanzee X divergence (1.56%)⁶. In any case, α is underestimated if Fig. 2a is used as the phylogeny.

As to the estimate of $\alpha = 2.1$ in ref. 7, corrections for multiple substitutions were not made in the data analysis. This could have underestimated α for old repetitive elements. For young elements, the low estimate of α could be in part due to low polymorphism on Y. One additional problem was the false assumption that the repetitive elements of the same subfamily were inserted into the genome at the same time¹⁶, introducing errors into the estimation of α .

We have provided a resolution for the controversy on the magnitude of α in primates. Most previous studies compared X- and Y-linked sequences (see ref. 2), and some have argued that the high α -values could be due to a reduction in mutation rate in X rather than an elevated rate in Y¹⁷. To avoid this possibility, we compared an autosomal and a Y-linked locus. The fact that this study and all previous studies, with the exception of comparisons between closely related species, give consistent estimates of $\alpha \approx 4$ –6 provides conclusive evidence for strong male-driven evolution in hominoids. □

Methods

DNA sequencing

We studied three homologous fragments on Y and chromosome 3 that correspond to nucleotides 36,887–40,238, 49,037–54,082, and 55,171–57,328 of human Y contig

AC006983.4 and nucleotides 59,378–56,044, 47,226–42,245, and 40,284–38,159 of human chromosome 3 contig AC010139.4. The 5' end of each fragment is located 1.5, 13.7 and 20.6 kb, respectively, downstream from the 3' untranslated region of the *DAZ* gene and they have no homology with an expressed sequence tags and contain no known or predicted gene. Interspersed and simple repeats were identified using RepeatMasker (<http://ftp.genome.washington.edu/RM/RepeatMasker.html>). Searches employing GenScan and BLAST failed to identify any genes or exons within these fragments.

Genomic DNAs of a female and a male human, bonobo (*Pan paniscus*), gorilla (*Gorilla gorilla*), gibbon (*Hylobates lar*) and siamang (*Hylobates syndactylus*) were used separately as templates in polymerase chain reaction (PCR). The Expand High Fidelity PCR system (Roche) was used to minimize errors in PCR amplification. The PCR primers were designed from the human sequence (available as Supplementary Information). Female DNA was used to amplify the chromosome 3 locus and male DNA to amplify the chromosome Y locus. The sequences were deposited in GenBank under accession numbers AF483550–AF483579.

Statistical analyses

The sequences of the fragments studied were concatenated and aligned using the MegAlign module of DNASTar (Lasergene), and the alignment was adjusted manually. As the divergence between the sequences studied is low (at most 11%), the alignment was rather simple. We found evidence of at least two copies of the *DAZ* locus on chromosome Y in human (there were five 'polymorphic' sites in a 10.4-kb sequence), gorilla (two polymorphic sites), and siamang (four polymorphic sites). Thus, the copies were almost identical, suggesting that the duplications happened recently.

To take the presence of polymorphisms on chromosome 3 into account, two pseudosequences were generated for each of the individuals studied. For example, if the site was heterozygous for A and G, one of the pseudosequences was assigned 'A' and the other was assigned 'G'. This assignment was done randomly by DAMBE¹⁸. Insertion/deletion polymorphisms (indels) in the alignment, as well as heterozygous indel sites, were not included in the calculations. These indels included one deletion of 106 base pairs (bp) in gibbon Y and one insertion of 33 bp in human Y, although most of the others were short indels, many of which were in mononucleotide microsatellites. One segment of 121 bp in gorilla chromosome 3 was difficult to sequence and was excluded from analysis. So, in total, 1,148 nucleotide sites were excluded from analysis. Tajima and Nei's genetic distances¹⁹ were estimated using DAMBE¹⁸. The distances were apportioned among the branches of the phylogeny using the FITCH module of PHYLIP²⁰. The ratio of substitution rates (Y/A) was calculated from the sum of the branches on Y (Y) and chromosome 3 (A). The average of the two values for the chromosome 3 pseudosequences was used in the calculations. This procedure was also used for treating the 'polymorphic' sites in Y-linked sequences. The formula from ref. 12, $Y/A = 2\alpha/(1 + \alpha)$, was used to calculate α . To calculate the 95% CI for α , we used two methods. First, we derived the variance of Y/A . If L is the length of the sequence, $V(Y) = Y(1 - Y)/[L(1 - 4Y/3)^2]$, $V(A) = A(1 - A)/[L(1 - 4A/3)^2]$, and $V(Y/A) = V(Y)/E(A)^2 + E(Y)^2V(A)/E(A)^4$. Second, we used the bootstrap. The results for the two methods were similar. The second method was used in the text.

Received 8 October 2001; accepted 25 January 2002.

1. Crow, J. F. The origins, patterns and implications of human spontaneous mutation. *Nature Rev. Genet.* **1**, 40–47 (2000).
2. Huang, W., Chang, B. H.-J., Gu, X., Hewett-Emmett, D. & Li, W.-H. Sex differences in mutation rate in higher primates estimated from AMG intron sequences. *J. Mol. Evol.* **44**, 463–465 (1997).
3. Agulnik, A. I. et al. Evolution of the *DAZ* gene family suggests that Y-linked *DAZ* plays little, or a limited, role in spermatogenesis but underlines a recent African origin for human populations. *Hum. Mol. Genet.* **7**, 1371–1377 (1998).
4. Nachman, M. W. & Crowell, S. L. Estimate of the mutation rate per nucleotide in humans. *Genetics* **156**, 297–304 (2000).
5. Ellegren, H. Heterogeneous mutation processes in human microsatellite DNA sequences. *Nature Genet.* **24**, 400–402 (2000).
6. Bohossian, H. B., Skaletsky, H. & Page, D. C. Unexpectedly similar rates of nucleotide substitution found in male and female hominids. *Nature* **406**, 622–625 (2000).
7. International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2000).
8. Pecon Slattery, J. & O'Brien, S. J. Patterns of Y and X chromosome DNA sequence divergence during the Felidae radiation. *Genetics* **148**, 1245–1255 (1998).
9. Ellegren, H. & Fridolfsson, A. K. Male-driven evolution of DNA sequences in birds. *Nature Genet.* **17**, 182–184 (1997).
10. Shimmin, L. C., Chang, B. H.-J., Hewett-Emmett, D. & Li, W.-H. Potential problems in estimating the male-to-female mutation rate ratio from DNA sequence data. *J. Mol. Evol.* **37**, 160–166 (1993).
11. Saxena, R. et al. Four *DAZ* genes in two clusters found in the *AZF* region of the human Y chromosome. *Genomics* **67**, 256–267 (2000).
12. Miyata, T., Hayashida, H., Kuma, K., Mitsuyasu, K. & Yasunaga, T. Male-driven molecular evolution: a model and nucleotide sequence analysis. *Cold Spring Harbor Symp. Quant. Biol.* **52**, 863–867 (1987).
13. Begun, D. J. & Aquadro, C. F. Levels of naturally occurring DNA polymorphism correlate with recombination rates in *D. melanogaster*. *Nature* **356**, 519–520 (1992).
14. Charlesworth, B. & Charlesworth, D. The degeneration of Y chromosomes. *Phil. Trans. R. Soc. Lond. B* **355**, 1563–1572 (2000).
15. Li, W.-H. Distribution of nucleotide differences between two randomly chosen cistrons in a finite population. *Genetics* **85**, 331–337 (1977).
16. Erlandsson, R., Wilson, J. F. & Paabo, S. Sex chromosomal transposable element accumulation and male-driven substitutional evolution in humans. *Mol. Biol. Evol.* **17**, 804–812 (2000).
17. McVean, G. T. & Hurst, L. D. Evidence for a selectively favourable reduction in the mutation rate of the X chromosome. *Nature* **386**, 388–392 (1997).

18. Xia, X. *Data Analysis in Molecular Biology and Evolution* (Kluwer Academic, Boston, 2000).
 19. Tajima, F. & Nei, M. Estimation of evolutionary distance between nucleotide sequences. *Mol. Biol. Evol.* **1**, 269–285 (1984).
 20. Felsenstein, J. PHYLIP—Phylogeny Inference Package (version 3.2). *Cladistics* **5**, 164–166 (1989).

Supplementary Information accompanies the paper on Nature's website (<http://www.nature.com>).

Acknowledgements

The DNA samples were purchased from San Diego Zoological Society and the gibbon sample was given by M. Jensen-Seaman. We thank J. Crow and D. Page for comments. This study was supported by NIH grants.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to W.-H.L. (e-mail: whli@uchicago.edu).

Future projections for Mexican faunas under global climate change scenarios

A. Townsend Peterson*, Miguel A. Ortega-Huerta†, Jeremy Bartley‡, Victor Sánchez-Cordero§, Jorge Soberón||, Robert H. Buddemeier‡ & David R. B. Stockwell¶

* Natural History Museum; † Department of Geography and Kansas Applied Remote Sensing Program; ‡ Kansas Geological Survey and Department of Geography, The University of Kansas, Lawrence, Kansas 66045, USA
 § Departamento de Zoología, Instituto de Biología, Apartado Postal 70-153;
 || Departamento de Ecología Evolutiva, Instituto de Ecología, Universidad Nacional Autónoma de México, D.F. 04510, México
 ¶ San Diego Supercomputer Center, University of California, 9500 Gilman Drive, La Jolla, California 92093, USA

Global climates are changing rapidly, with unexpected consequences¹. Because elements of biodiversity respond intimately to climate as an important driving force of distributional limitation², distributional shifts and biodiversity losses are expected^{3,4}. Nevertheless, in spite of modelling efforts focused on single species² or entire ecosystems⁵, a few preliminary surveys of fauna-wide effects^{6,7}, and evidence of climate change-mediated shifts in several species^{8,9}, the likely effects of climate change on species' distributions remain little known, and fauna-wide or community-level effects are almost completely unexplored⁶. Here, using a genetic algorithm and museum specimen occurrence data, we develop ecological niche models for 1,870 species occurring in Mexico and project them onto two climate surfaces modelled for 2055. Although extinctions and drastic range reductions are predicted to be relatively few, species turnover in some local communities is predicted to be high (>40% of species), suggesting that severe ecological perturbations may result.

We present a fauna-wide suite of predictions of the biodiversity consequences of global climate change. Taking advantage of the enormous biodiversity data resources accumulated by Mexico's Comisión para el Conocimiento y Uso de la Biodiversidad (CONABIO), as well as new tools in biodiversity informatics and quantitative geography¹⁰, we develop predictions of the effects of global climate change on distributions of 1,870 species (all 1,179 birds, all 416 mammals and all 175 butterflies of the families Papilionidae and Pieridae in Mexico) under two scenarios of global climate change (one conservative, one liberal) and three assump-

tions about dispersal ability¹¹. We use the Genetic Algorithm for Rule-set Prediction (GARP), a machine-learning system that has shown excellent predictive ability in delineating species' ecological niches and predicting geographic distributions¹². This first large-scale application of species-by-species models to the challenge of understanding the biodiversity consequences of global climate change provides a first view of the consequences of climate-change processes across a biodiversity-rich region.

To provide a detailed example of model results for a species, we examine the distribution of the west Mexican chachalaca (*Ortalis poliocephala*), a bird species endemic to tropical southwestern Mexico. Its present geographic distribution was summarized well by the ecological niche model, with correct prediction of 90% of independent test points, indicating that the model of the species' ecological niche was adequate. Under both scenarios of climate change, although the coastal portion of the species' distribution remained intact, the interior portion became less habitable, and a narrow band in the foothills of the coastal mountain ranges (Sierra Madre del Sur) more habitable (Fig. 1). Assuming universal dispersal abilities (not a good assumption), the species would encounter overall somewhat less (22.7% decline under the conservative climate change scenario) or more (75.1% increase under the liberal scenario) habitable area. However, under more realistic limited dispersal assumptions (dispersal into contiguous areas or no dispersal), the two scenarios agree more closely (conservative scenario 33.7% decrease, liberal scenario 29.7% decrease).

Examining patterns of change in distributional area across all 1,870 species, two climate-change scenarios, and three distributional assumptions analysed, consistent patterns emerge: species'

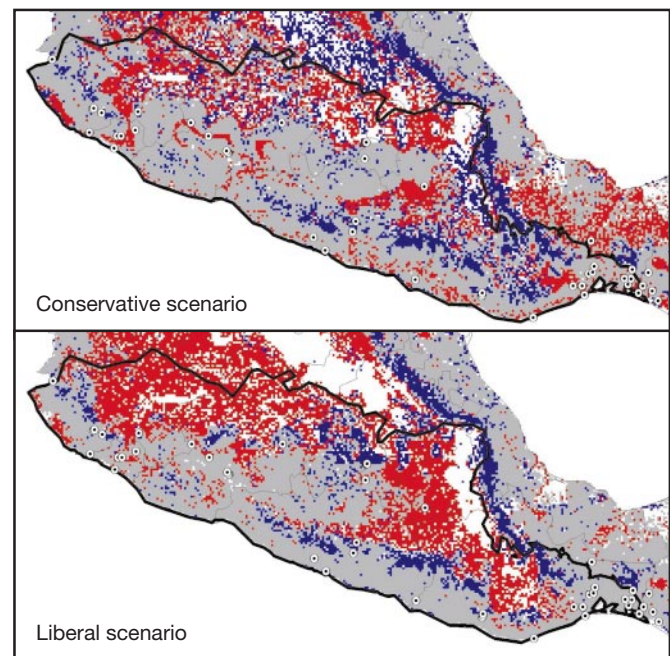


Figure 1 Example of analyses of effects of global climate change on a species' (*Ortalis poliocephala*) potential geographic distribution. Top, changes expected under the conservative scenario of the Hadley simulation; bottom, changes expected under the liberal scenario. Grey areas indicate conditions appropriate for the species both at present and under the scenario of change; white areas indicate conditions inappropriate for the species both at present and under the scenario of change; red areas are at present appropriate, but are predicted to become uninhabitable by the species; blue areas are at present not appropriate, but are predicted to become appropriate for the species; dotted circles represent known occurrence points; the black line indicates present geographic limits to species' distribution. Areas outside the black line correspond to the distributions of other *Ortalis* species, and represent areas of overprediction owing to historical factors²¹.

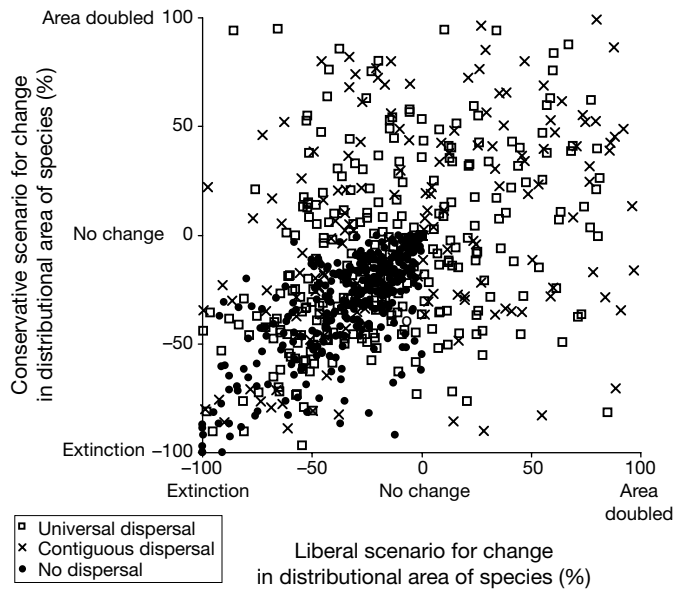


Figure 2 Distribution of expected percentage change in distributional area of 334 species endemic to Mexico under two scenarios of climate change (conservative and liberal) and three scenarios of dispersal abilities (universal, contiguous and no dispersal). -100%, extinction; 0%, no change; 100%, doubling of distributional area.

distributions expand and contract with equal frequency under the universal dispersal ability assumption, but show increasing tendencies to contract under the contiguous dispersal and no dispersal assumptions. To estimate the magnitude of distributional area changes to be expected without the biases introduced by political boundaries, we focused on 334 species endemic to Mexico (Fig. 2), none of which is predicted to expand its range outside of Mexico. Under the unrealistic assumption of universal dispersal ability, the species' distributional area expand and contract with equal frequency; results are similar under the contiguous dispersal assumption.

tion. Under the assumption of no dispersal ability, however, almost all species decline, and numbers of species suffering extreme reductions in distributional area (90% or more) are higher (2.4% of species). Reduction of distributional area is a consistently good predictor of species extinction, so these numbers represent minimum estimates of species in grave danger of extinction under these scenarios¹³: 0–2.4% of species are predicted to lose $\geq 90\%$ of the present distributional area, and 5.1–19.5% are predicted to lose $\geq 50\%$ of the present distributional area, under the three dispersal assumptions. The two climate-change scenarios are similar in the qualitative patterns predicted, although the liberal scenario predicts more dramatic effects in 11% of species (Fig. 2). No significant differences exist between taxonomic groups (mammals, birds, butterflies) in expected severity of effects on species' distributional areas; however, the butterfly families treated herein have relatively good dispersal abilities, and thus may not differ much from birds in this regard.

Gain and loss of species is distributed unevenly across Mexico (Fig. 3). Under the assumption of dispersal only into contiguous areas (averaging across the two climate-change scenarios), colonizations concentrate along the major sierras of Mexico, whereas extinctions are focused in the broad, open Chihuahuan desert and northwestern coastal plain. When we sum colonizations and extinctions for each grid cell (averaging between climate-change scenarios) across all species, and relate them to present species richness as a measure of species turnover, foci of turnover are evident in northern Mexico in the Chihuahuan desert, in interior valleys extending south to Oaxaca, and in the Baja California peninsula (predicted turnover rates as high as 45%). Because of the general northward trend of species' distributional shifts, and in view of the complicating effects of political boundaries, the southern quarter of the maps in Fig. 3 should be interpreted with caution: species moving into Mexico from the south, or incompletely modelled owing to Central American populations being excluded from our modelling efforts, may change the results.

The analyses presented herein illustrate the magnitude of effects that may be expected as a consequence of global climate change. Although only limited numbers of species will face entirely unsuit-

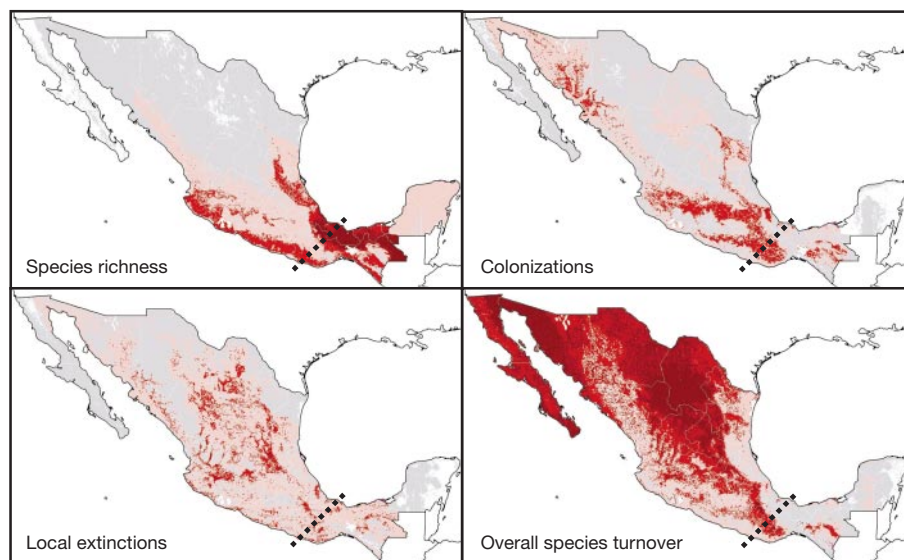


Figure 3 Modelled species turnover in biological communities (1,870 species) across Mexico. Modelled current species richness: white, <155 species; grey, 155–306 species; pink, 307–458 species; red, 459–610 species; dark red, 611–763 species. Local extirpations: white, <29 species; grey, 29–56 species; pink, 57–84 species; red, 85–112 species; dark red, 113–140 species. Colonizations: white, <25 species; grey,

25–48 species; pink, 49–71 species; red, 72–95 species; dark red, 96–119 species. Species turnover: white, <10%; grey, 10–20%; pink, 20–30%; red, 30–40%; dark red, >40. The southern quarter of these maps (indicated by dashed line), however, may be subject to some bias (see text), and thus should be interpreted with caution.

able conditions for persistence, others will experience drastic reductions and fragmentation of distributional areas, or extend their distributions, creating new natural communities with unknown properties. Indeed, some workers¹⁴ have argued that these reorganizations of communities will produce stronger distributional effects than will the direct effects of climate change on species' distributions; in this sense, our models provide null hypotheses of distributional expectations in the absence of strong interaction effects. More generally, these approaches to modelling distributional change caused by global climate change can be used to produce synthetic models of the combined effects of climate change and other scenarios (for example, species' invasions¹⁵), allowing us to address questions of productivity in economically important ecosystems, emerging diseases and other complex challenges¹⁶. □

Methods

The general approach to modelling climate change effects on biodiversity is developed in detail elsewhere¹¹, as are the details of the algorithm used for modelling species' ecological niches^{17,18}.

Data on distributions and ecological dimensions

Distributional data representing 112,456 records (that is, unique species × latitude–longitude combinations) for all 1,179 bird species in Mexico (a revised taxonomic treatment¹⁹; migrants and residents included), all 416 mammal species in Mexico, and 175 butterfly species (all the Papilionidae and Pieridae in Mexico) were obtained from continuing efforts to assemble distributional information on Mexican species from natural history museums around the world by the Comisión Nacional para el Uso y Conocimiento de la Biodiversidad (CONABIO; <http://www.conabio.gob.mx/>). Forty-five museums^{20–22} contributed data. Environmental data were provided on a 4 × 4 km grid by CONABIO (annual mean temperature in 15 classes, annual mean precipitation in 19 classes) and the Defense Mapping Agency (digital elevation model; <http://edcdaac.usgs.gov/gtopo30/hydro/namerica.html>). From the latter, coverages summarizing elevation, slope, aspect and potential solar radiation (all continuous) were extracted. The categorical temperature and precipitation maps were converted into pseudo-continuous maps using pycnophylactic interpolation, a mass-preserving spatial interpolation for categorical data²³.

Scenarios of climate change

The general circulation model used²⁴ (HadCM2) includes several scenarios. We assessed both a conservative and a less conservative view of how climates could change over the next 50 years using the HHGSDX50 and HHGGAX50 scenarios (http://ipcc-ddc.cru.uea.ac.uk/cru_data/examine/HadCM2_info.html). The HHGSDX50 scenario assumes an increase of 0.5% CO₂ per year (IS92d), and incorporates sulphate aerosol forcing, making it a relatively conservative estimate of climate change. The HHGGAX50 scenario assumes a 1% increase of CO₂ per year (IS92a), and does not allow for the effects of sulphate aerosols, and so is more liberal. Results are based on a 30-year average around 2055 (2040–2069); our models therefore do not take into account potential effects of increased climate variability (El Niño events, in particular) on species' distributions. Data are provided at a spatial resolution of 2.5° × 3.75°. Expected changes in temperature (°C) and precipitation (mm) under each scenario were extracted; to increase resolution, we interpolated to 0.5 × 0.5° cells using a nearest-neighbour contouring algorithm, as suggested in the data guidelines of the Intergovernmental Panel on Climate Change (<http://ipcc-ddc.cru.uea.ac.uk/>). These coverages of expected change were used to increment the pseudo-continuous current temperature and precipitation maps. The resulting estimated climate (annual averages) coverages of Mexico for the mid-twenty-first century indicate that the average annual mean temperature of Mexico may increase by 1.6–2.5°, and mean annual precipitation may decrease by 70–130 mm. Although inclusion of intermediate time-slice scenarios in such modelling efforts would be greatly desirable, in the present example, it is not necessary, as shifts in species' distributions were relatively subtle.

Ecological niche modelling and dispersal assumptions

The ecological niche of a species can be defined as the conjunction of ecological conditions within which it can maintain populations without immigration^{25,26}; as such, it is defined in multidimensional ecological/environmental space²⁷. Several approaches have been used to approximate species' ecological niches²⁸, of these, the most robust appears to be the genetic algorithm for rule-set prediction (GARP), which includes several inferential approaches in an iterative, artificial-intelligence-based approach¹⁷, and has had extensive testing^{12,17,29}. GARP works in an iterative process of rule selection, evaluation, testing, and incorporation or rejection to produce a heterogeneous (for example, logistic regression, bioclimatic rules) rule set describing the species' ecological niche.

Ecological niche models developed with GARP can be projected onto current and modelled future landscapes. Projection onto the current landscape indicates present-day geographic distribution of suitable conditions. Because species' distributions are limited by both ecological and historical factors (such as barriers to dispersal)²¹, we restricted predicted distributions to those ecoregions (<http://www.conabio.gob.mx/sig/>

[acerca_sig_pr.html](#)) in which the species had actually been recorded¹¹. GARP models can also be applied to climate-change-transformed landscapes to identify areas of potential distribution for a species after modelled environmental changes.

We synthesized pre-change maps and liberal and conservative post-change maps (averaged for extinction calculations) for each species by measuring potential distributional area under each of three assumptions regarding dispersal ability. An unrealistic assumption was that species could disperse to any site at which conditions were favourable for population persistence ('universal dispersal'), that is, raw, uncut distributions were compared before and after change. More realistic was the assumption that species would be able to disperse through continuous habitat but not jump over barriers ('contiguous dispersal'), that is, the modelled actual distribution was overlapped with the post-change prediction, and areas of continuous habitable environments that touch the present distribution were identified; discontinuities of one pixel or greater (~4 km) were considered barriers to dispersal. Another more realistic assumption was that species were simply unable to disperse and would inhabit only those portions of present distributional areas that remain habitable ('no dispersal'), that is, the modelled actual distribution was reduced to those areas predicted to be habitable post-change. These analyses assume no evolution in niche characteristics²¹, and do not take into account interactions among species such as competition, predation, and so on.

Community turnover

Each species was predicted to experience local extinction and colonization, depending on the climate-change scenario and dispersal assumptions. To evaluate geographic patterns in these shifts, under the contiguous areas dispersal assumption, we summed local extirpations (*E*) and local colonizations (*C*) separately across all species, and averaged the summed maps under the two climate-change scenarios. In this way, we summed distributional shifts across the entire fauna. We added present distributions across all species to create a map of current species richness (*S*), and calculated expected percentage species turnover as $100(E + C)/(S + C)$.

Received 15 June 2001; accepted 3 January 2002.

1. Houghton, J. T. et al. (eds) *IPCC Third Assessment Report: Climate Change 2001* (Cambridge Univ. Press, Cambridge, 2001).
2. Johnston, K. M. & Schmitz, O. J. Wildlife and climate change: Assessing the sensitivity of selected species to simulated doubling of atmospheric CO₂. *Glob. Change Biol.* **3**, 531–544 (1997).
3. Dobson, A., Jolly, A. & Rubenstein, D. The greenhouse effect and biological diversity. *Trends Ecol. Evol.* **4**, 64–68 (1989).
4. Chapin, F. S. I. Consequences of changing biodiversity. *Nature* **405**, 234–242 (2000).
5. Woodward, F. I., Lomas, M. R. & Betts, R. A. Vegetation-climate feedback in a greenhouse world. *Phil. Trans. R. Soc. Lond. B* **353**, 29–39 (1998).
6. Sala, O. E. et al. Global biodiversity scenarios for the year 2100. *Science* **287**, 1770–1773 (2000).
7. Price, J. Modeling the potential impacts of climate change on the summer distributions of Massachusetts passerines. *Birds Obs.* **28**, 224–230 (2000).
8. Visser, M. E., van Noordwijk, A. J., Tinbergen, J. M. & Lessells, C. M. Warmer springs lead to mistimed reproduction in great tits (*Parus major*). *Proc. R. Soc. Lond. B* **265**, 1867–1870 (1998).
9. Parmesan, C. Climate and species' range. *Nature* **382**, 765–766 (1996).
10. Soberon, J. Linking biodiversity information sources. *Trends Ecol. Evol.* **14**, 291 (1999).
11. Peterson, A. T. et al. Effects of global climate change on geographic distributions of Mexican Cracidae. *Ecol. Model.* **144**, 21–30 (2001).
12. Peterson, A. T. Predicting species' geographic distributions based on ecological niche modeling. *Condor* **103**, 599–605 (2001).
13. Koopowitz, H., Thornhill, A. D. & Andersen, M. A general stochastic model for the prediction of biodiversity losses based on habitat conversion. *Conserv. Biol.* **8**, 452–438 (1994).
14. Davis, A. J., Jenkinson, L. S., Lawton, J. H., Shorrocks, B. & Wood, S. Making mistakes when predicting shifts in species range in response to global warming. *Nature* **391**, 783–786 (1998).
15. Peterson, A. T. & Vieglais, D. A. Predicting species invasions using ecological niche modeling. *BioScience* **51**, 363–371 (2001).
16. Sánchez-Cordero, V. & Martínez-Meyer, E. Museum specimen data predict crop damage by tropical rodents. *Proc. Natl Acad. Sci. USA* **97**, 7074–7077 (2000).
17. Peterson, A. T., Stockwell, D. R. B. & Kluza, D. A. in *Predicting Species Occurrences: Issues of Scale and Accuracy* (ed. Scott, J. M.) (Island, Washington DC, in the press).
18. Stockwell, D. R. B. & Peters, D. P. The GARP modelling system: Problems and solutions to automated spatial prediction. *Int. J. Geograph. Inform. Systems* **13**, 143–158 (1999).
19. Peterson, A. T. & Navarro-Sigüenza, A. G. Alternate species concepts as bases for determining priority conservation areas. *Conserv. Biol.* **13**, 427–431 (1999).
20. Peterson, A. T., Navarro-Sigüenza, A. G. & Benítez-Díaz, H. The need for continued scientific collecting: A geographic analysis of Mexican bird specimens. *Ibis* **140**, 288–294 (1998).
21. Peterson, A. T., Soberón, J. & Sánchez-Cordero, V. Conservatism of ecological niches in evolutionary time. *Science* **285**, 1265–1267 (1999).
22. Llorente-Bousquets, J., Luis, A. & Vargas, I. *Papilionidae y Pieridae de México: Distribución Geográfica e Ilustración* (Universidad Nacional Autónoma de México, Mexico City, 1997).
23. Tobler, W. R. Smooth pycnophylactic interpolation of geographic regions. *J. Am. Stat. Ass.* **74**, 519–530 (1979).
24. Carson, D. J. Climate modelling: achievements and prospects. *Q. J. R. Meteorol. Soc.* **125**, 1–27 (1999).
25. Holt, R. D. & Gaines, M. S. Analysis of adaptation in heterogeneous landscapes: implications for the evolution of fundamental niches. *Evol. Ecol.* **6**, 433–447 (1992).
26. Grinnell, J. Field tests of theories concerning distributional control. *Am. Nat.* **51**, 115–128 (1917).
27. MacArthur, R. *Geographical Ecology* (Princeton Univ. Press, Princeton, 1972).
28. Austin, M. P., Nicholls, A. O. & Margules, C. R. Measurement of the realized qualitative niche: Environmental niches of five *Eucalyptus* species. *Ecol. Monogr.* **60**, 161–177 (1990).
29. Peterson, A. T., Ball, L. G. & Cohoon, K. C. Predicting distributions of tropical birds. *Ibis* **144**, e27–e32 (2002).

Acknowledgements

We thank A. G. Navarro-Sigüenza, J. E. Llorente-Bousquets, A. M. Luis, I. Vargas and L. Oniate for assembling distributional data, C. Thomas for help and advice, and E. Martínez-Meyer, G. Jiménez-Casas, S. Egbert and K. P. Price for collaboration. This research was supported by the US National Science Foundation, and grants from CONACyT and DGAPA to J. Soberón and V. Sánchez-Cordero.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to A.T.P. (e-mail: town@ukans.edu).

Extraction of a weak climatic signal by an ecosystem

Arnold H. Taylor^{*†}, J. Icarus Allen^{*} & Paul A. Clark[‡]

^{*} Plymouth Marine Laboratory, Prospect Place, Plymouth PL1 3DH, UK

[†] Department of Mathematics and Statistics, University of Plymouth, Plymouth, UK

[‡] Department of Geography, University of Sussex, Brighton BN1 9SJ, UK

The complexity of ecosystems can cause subtle¹ and chaotic responses to changes in external forcing². Although ecosystems may not normally behave chaotically³, sensitivity to external influences associated with nonlinearity can lead to amplification of climatic signals. Strong correlations between an El Niño index and rainfall and maize yield in Zimbabwe have been demonstrated⁴; the correlation with maize yield was stronger than that

with rainfall. A second example is the 100,000-year ice-age cycle, which may arise from a weak cycle in radiation through its influence on the concentration of atmospheric CO₂ (ref. 5). Such integration of a weak climatic signal has yet to be demonstrated in a realistic theoretical system. Here we use a particular climatic phenomenon—the observed association between plankton populations around the UK and the position of the Gulf Stream^{6,7}—as a probe to demonstrate how a detailed marine ecosystem model extracts a weak signal that is spread across different meteorological variables. Biological systems may therefore respond to climatic signals other than those that dominate the driving variables.

Figure 1 illustrates this relationship with the latitude of the Gulf Stream (the GSNW index^{6,7}) at five locations in the North Sea over a period of three decades using zooplankton data collected by the Continuous Plankton Recorder Survey⁸, together with zooplankton data from a site off the Northumberland coast⁹. These are expressed as cumulative sums of the square of the difference between the zooplankton and GSNW index series after each series has been converted to a mean of zero and unit standard deviation¹⁰. Cumulative sums provide a sensitive test of whether the relationship is constant with time. When the variables are positively correlated the slope of the graph will be close to zero, rising to 2.0 if the variables are uncorrelated. In the central and southern North Sea abrupt changes in slope show that the relationship has not been constant with time. Figure 1 also shows data for a freshwater lake in the north of England, Esthwaite water¹¹, and similar results have been obtained in a neighbouring lake, Windermere¹². The probabilities that slopes such as these might arise by chance (see Methods section) are also shown in Fig. 1. Similar probabilities calculated using the North Atlantic Oscillation (NAO) index¹³ instead of the GSNW index were much higher and all were over 0.1, showing that the GSNW teleconnection is distinct from the effects of the NAO.

In order to determine how the GSNW signal finds its way into the plankton, time series of hourly surface heat fluxes calculated from

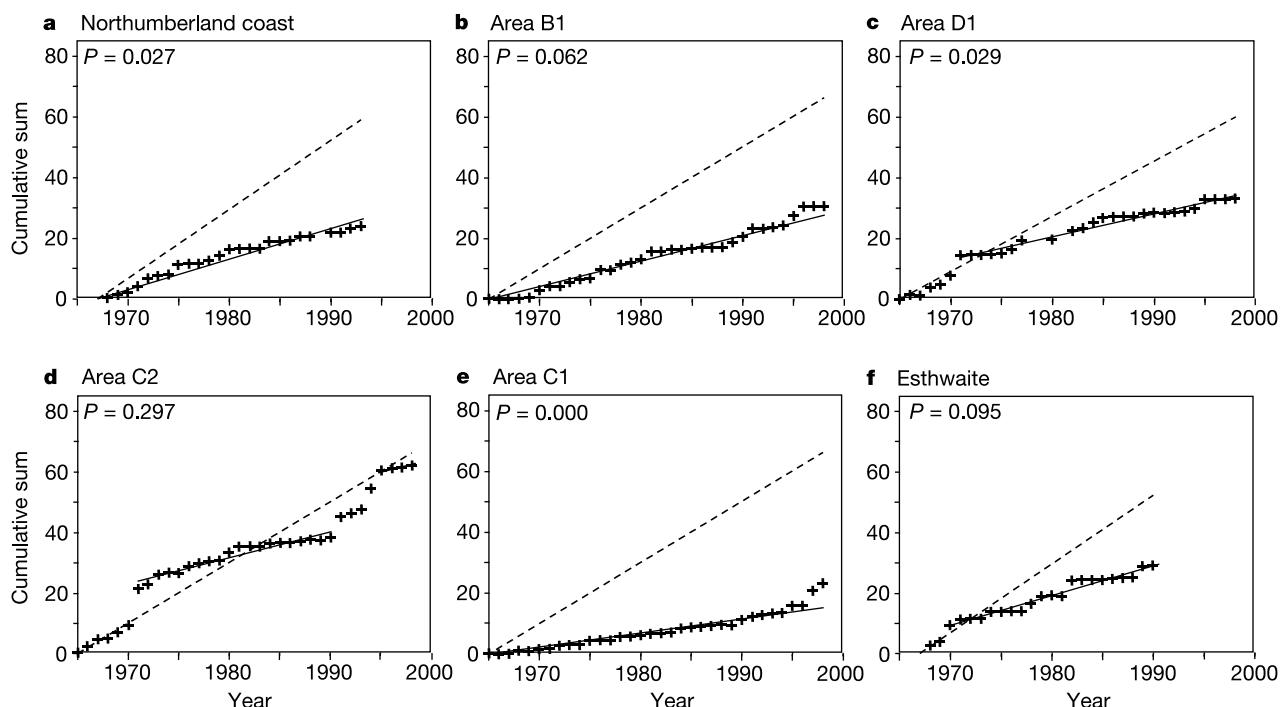


Figure 1 Correlation of observed plankton populations in and around the UK with the latitude of the Gulf Stream. Cumulative sum of the square of the difference between the abundance of total copepods and the GSNW index (crosses) at a site off the coast of Northumberland; in northern (B1), central (C1 and C2) and southern (D1) areas of the

North Sea and in Esthwaite water, UK (using *Daphnia* instead of total copepods). The solid lines are regression fits (included only to highlight the linear parts) and the broken lines show the results for an uncorrelated pair of time-series. The probabilities *P* that the slopes arose by chance are shown.

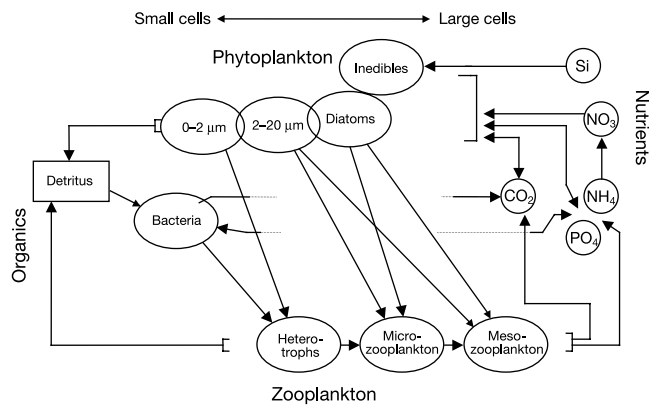


Figure 2 Trophic structure of the pelagic components of the ERSEM ecosystem model. See Methods for details.

meteorological observations made at Dublin from 1966 to 1998 were used to drive a coupled physical–ecological water-column model^{14,15}. This model is a synthesis of the European Regional Seas Ecosystem Model (ERSEM)¹⁶ with a one-dimensional version of the Princeton Ocean Model (POM)¹⁷. Details of the model are given in the Methods section and Fig. 2. The simulations are of a generic, seasonally stratified water column typical of the northern North Sea and as such may be representative of the areas in Fig. 1.

The solid lines in Fig. 3 show cumulative sum plots calculated for six of the variables from time series of April–September means produced by the model. Five of these graphs—the abundances of mesozooplankton (Fig. 3a), heteroflagellates (Fig. 3c) and total phytoplankton carbon (Fig. 3d), and the rates of total primary production (Fig. 3e) and bacteria uptake of inorganic nitrogen (Fig. 3f)—show a prolonged period through the 1970s and 1980s

when the slope was considerably less than the random value of 2, with graphs similar to those presented in Fig. 1. The probabilities of these occurring in random data was less than 0.05 in each case. This pattern is widespread in the ERSEM output. When the probabilities were calculated for 25 of the major standing stocks and fluxes, 19 were 0.1 or less. If the variables were independent such frequencies would only be expected to occur much less than once in a thousand times by chance. The variables in ERSEM are not, of course, independent. When principal components analysis was applied to the 25 time series it was found that 8 variables accounted for 98% of the total variance. If the degrees of freedom are reduced proportionally the frequencies would still occur by chance less than 1 in 20 times. As in the data of Fig. 1, these relationships are not just the effect of the NAO; using the NAO index instead of the GSNW index raises the 25 probabilities from an average of 0.11 to one of 0.32, even though the NAO is possibly the most important constituent of the meteorological observations¹³.

Analysis of the meteorological observations reveals that the GSNW signal is scattered through the data rather than localized in any particular subset. Cumulative sum plots and probabilities calculated from seasonal values of the main atmospheric driving variables—insolation and wind-strength—are shown in Fig. 4a and b. All of these graphs, apart from the wind strength in winter, have probabilities of well over 0.1, implying that the GSNW signal is only a weak constituent of these atmospheric data. These probabilities are much larger than those obtained from the ERSEM output. This relative weakness of the signal in the meteorological observations has been noted previously^{7,18}.

The way in which this weak signal comes to appear in the ERSEM output can be illustrated by comparing the model results in Fig. 3 with the other graphs on the figure in which either the insolation or the wind strength was fixed at its mean annual cycle. Four of the variables—total phytoplankton carbon (Fig. 3d), total primary production (Fig. 3e), bacterial nitrogen uptake (Fig. 3f) and

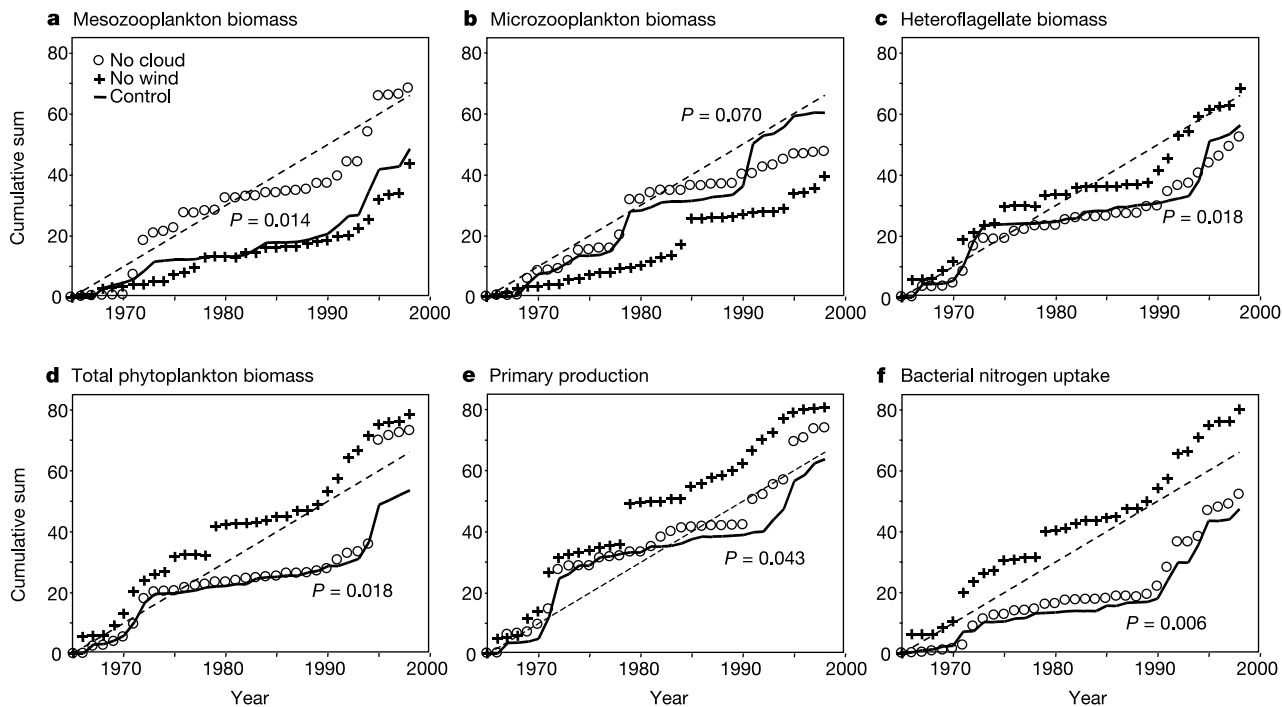


Figure 3 Correlation of simulated plankton populations with the latitude of the Gulf Stream. Cumulative sum of the square of the difference between the predicted time series of mean summer values (April to September) obtained from a run of the ERSEM model for 1966 to 1998 and the GSNW index. Plots (solid lines) are shown for the abundance of mesozooplankton, microzooplankton, heteroflagellates and all phytoplankton, the total

rate of primary production and the ammonium uptake by bacteria from the control run. The broken lines show the results for an uncorrelated pair of time-series. The circles and the crosses show the results when interannual variations were removed from the incident irradiance (no cloud) and the wind strength (no wind), respectively.

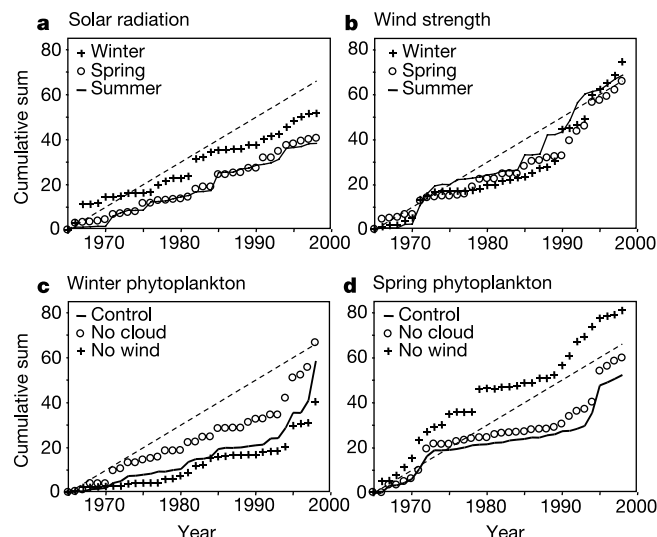


Figure 4 Correlation of meteorological forcing and phytoplankton populations with the latitude of the Gulf Stream. Cumulative sum of the square of the difference between the GSNW index and averages over winter (crosses), spring (circles) and summer (solid line) of solar irradiance (a) and wind strength (b), and averages over winter (c) and spring (d) of total phytoplankton abundance predicted by ERSEM. In c and d the solid line shows the standard model run, and the circles and crosses indicate those when interannual variations were removed from the incident irradiance and the wind strength, respectively. The broken lines show the results for an uncorrelated pair of time-series.

heterotrophic flagellate abundance (Fig. 3c)—show very little change when interannual variability is removed from the solar radiation, but have a random relationship with the GSNW index when the variability is removed from the wind strength. Over the summer period, means of these variables will be dominated by nutrient limitation when fluctuations in the supply of nutrients associated with changes in windiness are important. In contrast, the GSNW signal in the mesozooplankton is much more dependent on the radiation changes than on the wind variations and it is the non-varying insolation results that are close to the random line. Micro-zooplankton, which graze on the phytoplankton and heteroflagellates and which are preyed on by the mesozooplankton, show a more complex relationship resulting from both top-down and bottom-up control.

The processes operating in these calculations are revealed by Fig. 4c and d which show the same calculations carried out for the total phytoplankton abundance over the three months January to March and over April to June. When there is no year-to-year wind variability the GSNW index signal is removed from the spring to early summer period alone. This is a time when, because of density stratification, nutrient stress is important in limiting primary production, which will therefore be enhanced by wind-induced pulses of nutrients across the thermocline.

Removing variability from the cloud has the largest effect on phytoplankton abundance in the winter period but has little effect in the nutrient-limited spring–summer period. Figure 3 shows that the mesozooplankton are dependent on changes in solar irradiance and this implies that mesozooplankton abundance in summer is very dependent on phytoplankton abundance earlier in the year. This is consistent with Colebrook's observation¹⁹ that much of the interannual variability of zooplankton in the North Sea has its origins in winter, with its occurrence through the year being a function of persistence.

The critical process appears to be the control of phytoplankton growth by vertical turbulent mixing in winter, as discussed in ref. 20. Both spring–summer mesozooplankton and winter phytoplankton abundances are strongly correlated with the winter temperature difference between the surface and the bottom (cumulative sum

probabilities 0.002 and 0.000, respectively) even though the mean temperature difference is only 0.07 °C. Corresponding probabilities with the surface temperature are much larger at 0.41 and 0.26. The temperature difference is more correlated with the GSNW index (probability 0.12) than is the surface temperature (0.8).

Even so, there is still an additional effect from summer populations because the mesozooplankton graph, in the absence of wind variability (Fig. 3), shows less similarity to those of the primary producers and bacteria, not having such a clear plateau from 1972 to 1990. This may be part of the reason the mesozooplankton abundance is more strongly correlated with the GSNW index than is the winter temperature difference.

Although it remains unclear what mechanism beyond chance might be responsible for the associations between the atmospheric observations and the GSNW index, we have used the signal as a probe to show how a complex ecosystem can extract information that is scattered throughout a set of meteorological variables. A general consequence of our results is that biological systems can exhibit responses to subtle climatic signals, signals that may be distinct from those that are the most apparent. This could be a significant feature of future responses to climate change. A secondary consequence is the possibility that changes in biological populations could be sensitive indicators of previously unknown climatic processes. Our results also indicate a degree of predictability for summer zooplankton abundance on the basis of knowledge of winter conditions. □

Methods

The GSNW index calculated from the latitude of the north wall of the Gulf Stream close to the coast of the USA^{6,7}. The estimate of total copepods used as a measure of total zooplankton abundance mainly reflects the numerical abundance of smaller zooplankton taxa.

Probabilities of cumulative sums

The test of statistical significance applied to the graphs was designed to determine objectively the likelihood of low, flat regions occurring. It therefore needs to examine both how small is the slope and the length of the flat period, while taking account of serial correlation in the data. In the test used, 2,000 random plankton series were generated, each having the same serial correlation as the observed series. To measure the slope of each of the graphs, a slope was calculated between every pair of points (adjacent and non-adjacent), and these values were averaged. The slope will be less than that of the random line (2) for a positive correlation but can be greater than 2 if the correlation is negative. In order to focus on the probability of low slopes, slopes greater than 2 were excluded from the average. The average slope was then compared with those calculated by applying exactly the same procedure to each of the random plankton series. The number of random series giving a slope at least as small as the average gives an estimate of the probability (*P*) that such a small slope would arise by chance.

ERSEM model

The physical model simulates the vertical density structure and mixing processes throughout the water column. ERSEM describes the biogeochemical cycling of carbon, nitrogen, phosphorus and silicate through the pelagic food web, and has a structure (Fig. 2) which allows it to switch from a mesotrophic ecosystem to one dominated by the microbial loop. It is one of the most complete marine ecosystem models that have been developed.

ERSEM considers the ecosystem to be a series of interacting physical, chemical and biological processes that together exhibit coherent system behaviour. The dynamics of biological functional groups are described by population processes (growth, migration and mortality) and physiological processes (ingestion, respiration, excretion and egestion). The ecosystem is subdivided into three functional types: producers (phytoplankton), decomposers (pelagic bacteria) and consumers (zooplankton).

These broad functional classifications are then subdivided, by grouping biota according to their trophic level (subdivided according to size classes or feeding method) to create a food web (Fig. 2). Physiological processes and population dynamics are described by fluxes of carbon or nutrients between functional groups. It is assumed that whatever compositional changes occur within each pool over time, they are not large enough to cause substantial and persistent errors in the prediction of pool scale rate processes. State variables have been chosen in order to keep the model relatively simple without omitting any component that exerts a significant influence upon the energy balance of the system.

The chemical dynamics of nitrogen, phosphorus, silicate and oxygen are coupled to the biologically driven carbon dynamics. The combination of a food web with coupled nutrient dynamics allows the model to adjust to spatial and temporal variations in carbon and nutrient availability and reproduce the different types of ecosystem behaviour¹⁶.

The phytoplankton pool is described by three functional groups based on size and trophic position²¹: diatoms (20–200 µm), consumed by micro and mesozooplankton;

autotrophic flagellates (2–20 μm), consumed by micro and mesozooplankton; picoplankton (0.2–2 μm), consumed by heterotrophic nanoflagellates; and inedible phytoplankton $>20 \mu\text{m}$. The uptake of nutrients (NO_3 , NH_4 and PO_4) have been decoupled from the carbon assimilation processes by including dynamic nutrient kinetics²², whereby nutrient uptake is dependent on both the level of intercellular storage and external nutrient concentrations. The microbial food web contains bacteria, heterotrophic flagellates and mesozooplankton, each with dynamically varying C:N:P ratios and is described in ref. 23. Bacteria consume DOC, decompose detritus and can compete for inorganic nutrients with phytoplankton. Heterotrophic flagellates feed on bacteria and picoplankton and are consumed by microzooplankton and mesozooplankton. Microzooplankton feed on diatoms, autotrophic and heterotrophic flagellates and are consumed by mesozooplankton. Mesozooplankton feed on diatoms, autotrophic flagellates and microzooplankton²⁴. All three grazer groups are cannibalistic.

Simulations were made with the ERSEM parameter sets used in ref. 15, for the Humber plume region of the North Sea. The model was forced by heat fluxes calculated from meteorological data observed at Dublin. Although data for Dublin are not from the North Sea, Dublin lies directly in the path of weather systems which commonly move from the Gulf Stream to the North Sea.

Received 25 September 2001; accepted 24 January 2002.

1. Mysterud, A., Stenseth, N. C., Yoccoz, N. G., Langvæn, R. & Steinhel, G. Nonlinear effects of large-scale climatic variability on wild and domestic herbivores. *Nature* **410**, 1096–1099 (2001).
2. Sugihara, G., Grenfell, B. & May, R. M. Distinguishing error from chaos in ecological time series. *Phil. Trans. R. Soc. Lond. B* **330**, 235–251 (1990).
3. Berryman, A. A. & Millstein, J. A. Are ecological systems chaotic—and if not, why not? *Trends Ecol. Evol.* **4**(1), 26–28 (1989).
4. Cane, M. A., Eshel, G. & Buckland, R. W. Forecasting Zimbabwean maize yield using eastern equatorial Pacific sea surface temperature. *Nature* **370**, 204–205 (1994).
5. Shackleton, N. J. The 100,000-year ice-age cycle identified and found to lag temperature, carbon dioxide and orbital eccentricity. *Science* **289**, 1897–1902 (2000).
6. Taylor, A. H., Colebrook, J. M., Stephens, J. A. & Baker, N. G. Latitudinal displacements of the Gulf Stream and the abundance of plankton in the north-east Atlantic. *J. Mar. Biol. Assoc. UK* **72**, 919–921 (1992).
7. Taylor, A. H. North-south shifts of the Gulf Stream and their climatic connection with the abundance of zooplankton in the UK and its surrounding seas. *ICES J. Mar. Sci.* **52**, 711–721 (1995).
8. Glover, R. S. The continuous plankton recorder survey of the North Atlantic. *Symp. Zool. Soc. Lond.* **19**, 189–210 (1967).
9. Frid, C. L. J. & Hulselman, N. V. Far field control of long term changes in Northumberland (NW North Sea) coastal zooplankton. *ICES J. Mar. Sci.* **53**(6), 972–977 (1996).
10. Taylor, A. H. in *Changing States of Large Marine Ecosystems of the North Atlantic* (eds Sherman, K. & Skjoldal, H.-R.) (in the press).
11. George, D. G. The impact of regional-scale changes in the weather on long-term dynamics of *Eudaptomus* and *Daphnia* in Esthwaite Water, Cumbria. *Freshwat. Biol.* **45**, 111–121 (2000).
12. George, D. G. & Taylor, A. H. UK lake plankton and the Gulf Stream. *Nature* **378**, 139 (1995).
13. Hurrell, J. W. Decadal trends in the North Atlantic Oscillation: regional temperatures and precipitation. *Science* **269**, 676–679 (1995).
14. Allen, J. I., Blackford, J. C. & Radford, P. J. A 1-D vertically resolved modelling study of the ecosystem dynamics of the middle and southern Adriatic Sea. *J. Mar. Sys.* **18**, 265–286 (1998).
15. Allen, J. I., Howland, R. M. H., Bloomer, N. & Uncles, R. J. Simulating the spring phytoplankton bloom in the Humber Plume, UK. *Mar. Poll. Bull.* **37**, 295–305 (1999).
16. Baretta, J. W., Ebanhoh, W. & Ruardij, P. (eds) The European regional seas ecosystem model II. *J. Sea Res.* **38**(3/4), 229–483 (1997).
17. Blumberg, A. F. & Mellor, G. L. in *Mathematical Modelling of Estuarine Physics* (Proc. Int. Symp. Hamburg, 1978) (eds Sunderland & Holtz) 203–214 (Springer, Berlin, 1980).
18. Taylor, A. H. North-south shifts of the Gulf Stream: ocean-atmosphere interactions in the North Atlantic. *Int. J. Climatol.* **16**, 559–583 (1996).
19. Colebrook, J. M. Continuous plankton records: overwintering and annual fluctuations in the abundance of zooplankton. *Mar. Biol.* **84**, 261–265 (1985).
20. Huisman, J., van Oostven, P. & Weissing, F. J. Critical depth and critical turbulence: Two different mechanisms for the development of phytoplankton blooms. *Limnol. Oceanogr.* **44**(7), 1781–1787 (1999).
21. Ebenhö, W., Baretta, J. W. & Baretta-Bekker, J. G. The primary production module in a marine ecosystem model ERSEM II. *J. Sea Res.* **38**, 173–194 (1997).
22. Droop, M. R. The nutrient status of algal cells in continuous culture. *J. Mar. Biol. Assoc. UK* **54**, 825–855 (1974).
23. Baretta-Bekker, J. G., Baretta, J. W., Hansen, A. S. & Riemann, B. An improved model of carbon and nutrient dynamics in the microbial food web in marine enclosures. *Aquat. Microb. Ecol.* **14**, 91–108 (1998).
24. Broekhuizen, N. R., Heath, M. R., Hay, S. G. & Gurney, S. C. Modelling the dynamics of the North Sea's mesozooplankton. *Neth. J. Sea Res.* **33**, 381–406 (1995).

Acknowledgements

We wish to thank J. Stephens and J. Dearman for assisting with the calculations. B. Clarke provided statistical advice. A.H.T. is a Fellow of the Sir Alister Hardy Foundation for Ocean Science, which provided the plankton data. This work is part of the Core Strategic Programme of Plymouth Marine Laboratory.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to A.H.T. (e-mail: aht@pml.ac.uk).

Direct visuomotor transformations for reaching

Christopher A. Buneo, Murray R. Jarvis, Aaron P. Batista* & Richard A. Andersen

Division of Biology, California Institute of Technology, Mail Code 216-76, Pasadena, California 91125, USA

The posterior parietal cortex (PPC) is thought to have a function in the sensorimotor transformations that underlie visually guided reaching, as damage to the PPC can result in difficulty reaching to visual targets in the absence of specific visual or motor deficits¹. This function is supported by findings that PPC neurons in monkeys are modulated by the direction of hand movement, as well as by visual, eye position and limb position signals^{2–9}. The PPC could transform visual target locations from retinal coordinates to hand-centred coordinates by combining sensory signals in a serial manner to yield a body-centred representation of target location^{10–12}, and then subtracting the body-centred location of the hand. We report here that in dorsal area 5 of the PPC, remembered target locations are coded with respect to both the eye and hand. This suggests that the PPC transforms target locations directly between these two reference frames. Data obtained in the adjacent parietal reach region (PRR) indicate that this transformation may be achieved by vectorially subtracting hand location from target location, with both locations represented in eye-centred coordinates.

The problem that we address here is shown in Fig. 1a. Although the execution of movement requires the specification of a detailed pattern of inputs to the muscles, movement planning is believed to involve the computation of higher level movement parameters, such as the direction and/or distance that the hand must move to reach the target (vector M)¹⁰. This is due to the fact that movement goals, as well as evidence of our success in achieving these goals, are largely expressed in high level terms, that is, as visually perceived discrepancies between the position of the hand and target or deviations from a desired path¹³. Hereafter we use the term 'target position in hand coordinates' to describe vector M in Fig. 1a, although the terms 'movement vector' and 'motor error' could also be used. This information could be derived by subtracting the sensed location of the hand (vector H) from the sensed location of the target (vector T), as long as hand position and target position are coded in a common frame of reference. However, although target position appears to be coded in eye-centred (retinal) coordinates in the early stages of reach planning¹⁴, hand position is derived from both visual and proprioceptive signals, and can conceivably be coded in eye-centred coordinates, body centred coordinates (that is, with respect to the torso), or both. It is unclear therefore whether the operation shown in Fig. 1a is achieved by subtracting the position of the hand from the position of the target directly, using eye-centred coordinates (Fig. 1a, b), or by transforming target locations from eye- to head- to body-centred coordinates, and then subtracting the body-centred position of the hand^{10,11} (Fig. 1c).

We have approached this problem by analysing the reach-related activity of neurons in the PPC, while varying target position, hand position and gaze direction. Single cell recordings were obtained from area 5 (Fig. 2a, b), a subdivision of the PPC that projects directly to cortical and subcortical motor structures^{15–17}. In an initial experiment, 89 neurons from two monkeys were studied under four experimental conditions (Fig. 2c). In two conditions, gaze was held constant at the centre position of a vertically oriented

* Present address: Howard Hughes Medical Institute, and Department of Neurobiology, Stanford University School of Medicine, Fairchild Building, Room D209, Stanford, California 94305, USA

board of push buttons, and initial hand location was varied to the left (condition 1) or right (condition 2) of this centre position. In the other two conditions, initial hand location was located at the centre position and gaze was varied either to the left (condition 3) or right (condition 4) of centre. Across the four conditions, mean firing rate varied least when target locations were identical in both hand and eye coordinates. For example, the activity of the neuron in Fig. 2 varied substantially between conditions 1 and 2, when target locations were identical in hand coordinates but different in both body coordinates (with respect to the torso) and eye coordinates (with respect to the fixation point). The activity of this neuron also varied substantially between conditions 3 and 4, when target locations were identical in hand and body coordinates but different in eye coordinates. However, activity was very similar between conditions 1 and 4 and between conditions 2 and 3, when target locations were identical in both hand and eye coordinates.

This point is shown more generally in Fig. 3a–e. In each scatterplot, individual data points correspond to the mean firing rate of a single cell for two movements to the same target location in hand, body, eye, body and hand, or eye and hand coordinates (panels a–e, respectively). Data from all cells and all possible target locations are shown. The scatter in each plot illustrates how well that particular coordinate frame or frames accounted for area 5 neuronal population activity. A low degree of scatter (that is, a high degree of correlation) indicates a good fit to the data, whereas a high degree of scatter (low correlation) indicates a poor fit. Statistical analyses of the data in Fig. 3 revealed that area 5 activity was best correlated when target locations were identical in both eye and hand coordi-

nates (Kruskal–Wallis test with nonparametric multiple comparisons, $P < 0.05$).

Although the data in Fig. 3 could be explained by the presence of two populations of neurons, one coding in eye coordinates and the other in hand coordinates, Fig. 4 shows that the responses of neurons in area 5 are more consistent with a single population coding target location in both reference frames. Figure 4a–d shows the responses of an idealized neuron coding target location in eye coordinates (panels a, b), both eye and hand coordinates (panel c) or hand coordinates (panel d). Figure 4a shows that if target locations are coded purely in eye coordinates, tuning curves for target location in eye coordinates will not shift when initial hand location in eye coordinates is varied, although they will modulate in amplitude if initial hand location in eye coordinates is explicitly coded, as in Fig. 4b. Figure 4d shows that if target locations are coded purely in hand coordinates, tuning curves will shift completely with initial hand location. However, Fig. 4c shows that if target locations are coded in both eye and hand coordinates, tuning

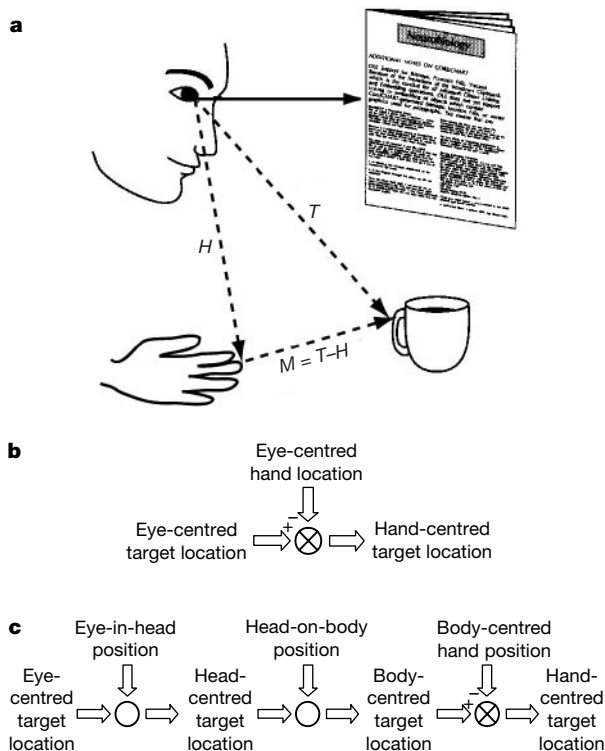


Figure 1 Visuomotor transformation schemes. **a**, Example of reaching for a cup while fixating on a newspaper. The position of the cup is initially represented in the brain in terms of its location on the peripheral retina (T). To reach for the cup, its position with respect to the hand must be known (M). This information could be acquired by directly subtracting hand position (H) from target position (T) in eye coordinates (**a**, **b**), or by gradually transforming the position of the target from eye- to body-centred coordinates, and subtracting the body-centred position of the hand (**c**). (Adapted from ref. 29.)

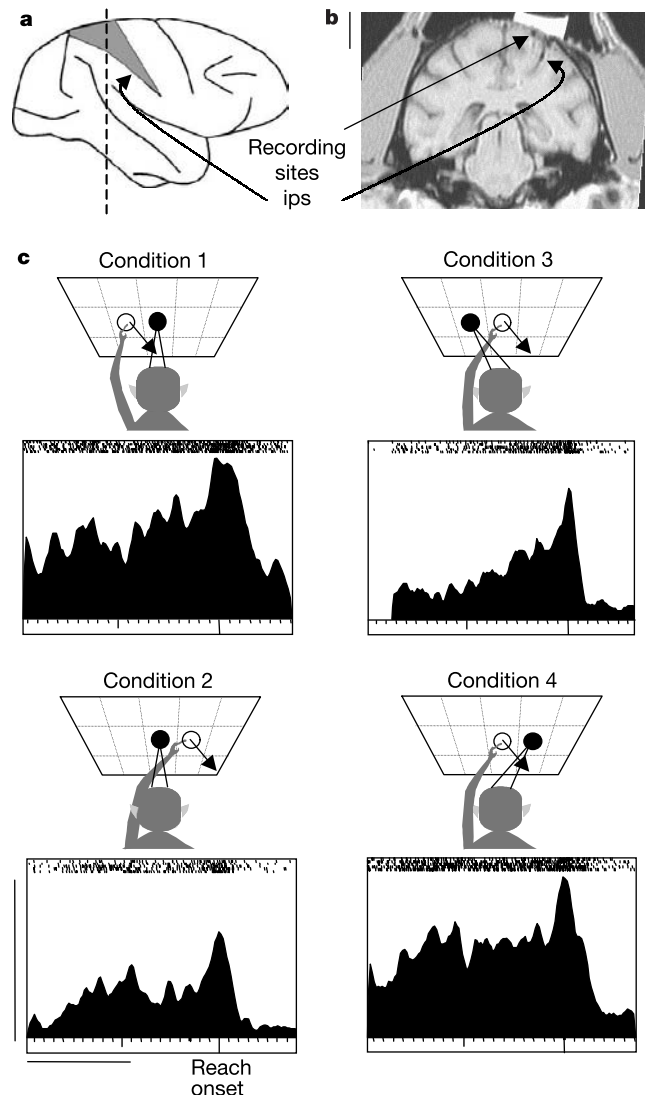


Figure 2 Responses of a single neuron from area 5. **a**, Diagram of a macaque monkey brain showing the location of area 5 (shaded region) and the approximate location of the magnetic resonance image (MRI) section in **b**. **b**, Coronal T1-weighted MRI section through the approximate centre of the area 5 recording sites. Scale bar, 1 cm. **c**, Spike density histograms of the activity of one area 5 neuron for the same planned movement vector (down and to the right) in each of four experimental conditions. Vertical scale bar, 140 spikes s^{-1} ; horizontal scale bar, 1 s.

curves for identical target locations in eye coordinates will partially shift with initial hand location, reflecting a 'compromise' between the eye and hand reference frames. Figure 4e shows for area 5 ($N = 89$) the distribution of tuning curve shifts when target locations in eye coordinates were the same and initial hand location was varied from left to right by 36° . This distribution has a single large peak at 18° (standard error ± 3), corresponding to a partial shift and consistent with a simultaneous coding of target location in both eye and hand coordinates. For comparison, Fig. 4e shows the same analysis applied to the PRR data ($N = 98$) of ref. 14. This distribution has a large peak at 0° , corresponding to no shift and consistent with a purely eye-centred coding of target location.

To explore further the hypothesis that neurons in area 5 code target location in both eye and hand coordinates, we trained one monkey to reach to a row of five targets from each of five starting locations, while maintaining fixation at a single board location (one position to the right of centre, see Fig. 5a). This design was chosen so

that responses of area 5 neurons could be compared directly with the responses predicted by the different coding models in Fig. 4. Figure 5b shows a contour plot of data obtained from one area 5 neuron in this experiment. These contours have a largely oblique orientation, similar to those of the idealized neuron in Fig. 4c. To facilitate comparisons between real response fields and their idealized counterparts, we summarized the trends giving rise to such contours by calculating the gradient of each response field and taking its resultant (see Methods for definition). The vector field superimposed on the contour plot in Fig. 5b shows the gradient that was obtained for this neuron. The resultant of the gradient in Fig. 5b is shown in Fig. 5c (blue vector) along with a 'population resultant' derived from 15 area 5 neurons (black vector), and the gradient resultant obtained for the idealized neuron in Fig. 4c (red vector). The resultants derived from experimental data have an orientation very similar to the one derived from the idealized responses, pointing largely down and to the right, which is consistent with a

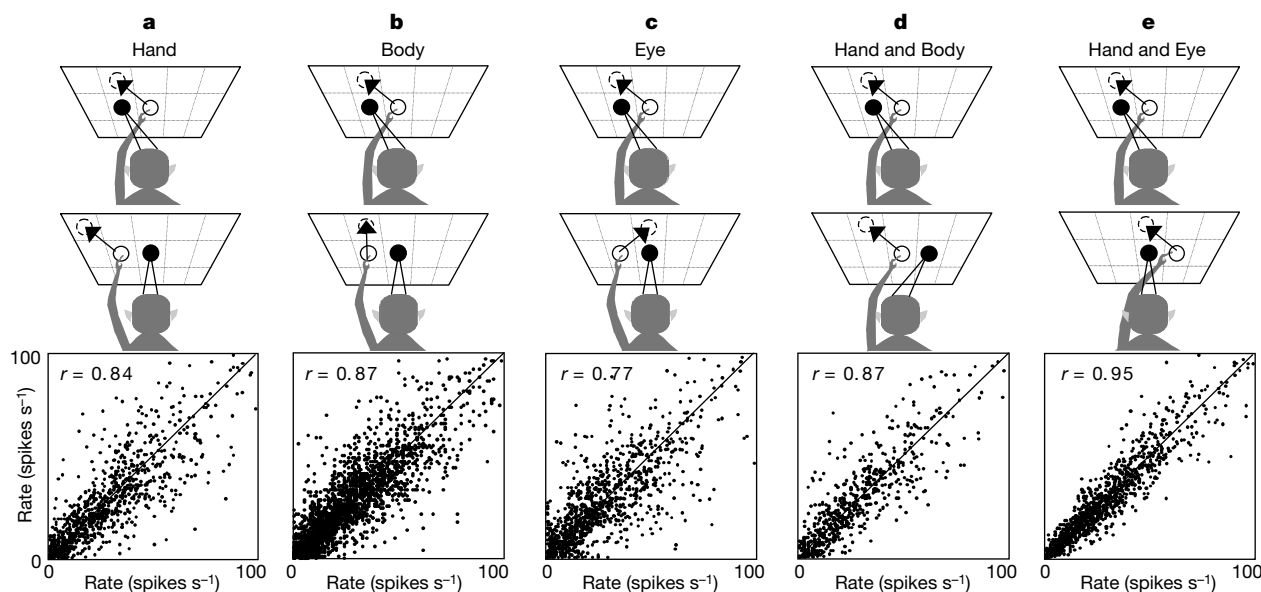


Figure 3 Area 5 neuronal population activity for reaches to identical target locations in hand coordinates (a), body coordinates (b), eye coordinates (c), hand and body coordinates (d), and hand and eye coordinates (e). Each data point corresponds to a cell's firing rate for a pair of movements; movements in a pair were taken from different

experimental conditions and were randomly assigned to the ordinate or abscissa. Example experimental conditions used to construct each figure (as well as an example pair of movements) are shown above the scatterplots.

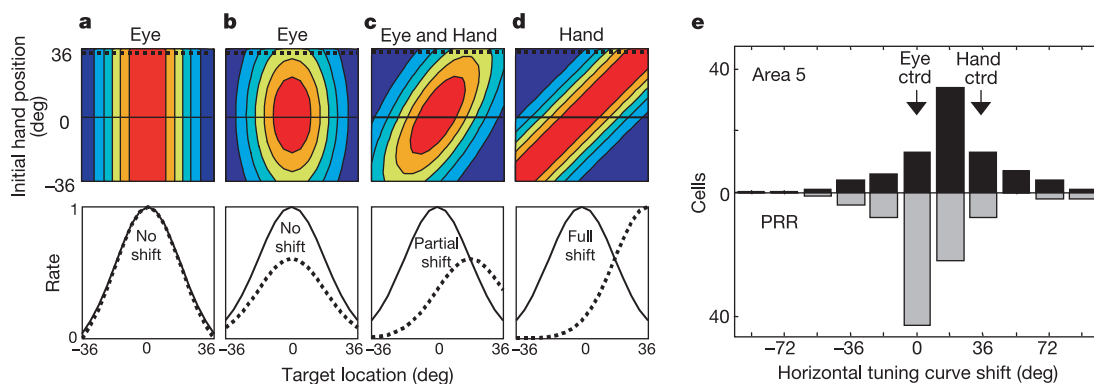


Figure 4 Shifting and non-shifting response fields in the PPC. a–c, Responses of an idealized neuron coding target location in eye coordinates (a), target location and initial hand location in eye coordinates (b), target location in eye and hand coordinates (c), and target location in hand coordinates (d). Responses are plotted as a function of horizontal

target location and horizontal initial hand location in eye coordinates. Tuning curves below each contour plot represent slices through each response field at initial hand locations of 0° (black line) and 36° (dotted line). e, Distribution of horizontal tuning curve 'shifts' for area 5 and PRR¹⁴.

coding of target location in both eye and hand coordinates. For comparison, Fig. 5d, e shows data obtained from PRR in the same experiment. The contours in Fig. 5d appear most similar to the idealized neuron in Fig. 4b, that is, they appear consistent with a coding of target location and initial hand location in eye coordinates. The gradient resultant for this neuron is shown in Fig. 5e (blue vector), along with a population resultant derived from 17 PRR neurons (black vector) and the resultant obtained for the idealized neuron in Fig. 4b (red vector). All three resultants point largely to the right in Fig. 5e, consistent with a coding of target location in eye coordinates.

The present results indicate that remembered target locations are coded in both eye and hand coordinates in area 5. These results are not unique to the perireach epoch—similar results were obtained when activity occurring immediately after presentation of the target, as well as during the delay period, was analysed separately. Thus, reaching to remembered target locations seems to entail a direct transformation from eye-centred to hand-centred coordinates, although more indirect schemes (for example, Fig. 1c) may still operate in other contexts¹⁸. The responses of neurons in the PRR suggest a mechanism for this direct transformation. Although PRR neurons strongly code target location in eye coordinates, some cells are 'gain modulated'¹⁹ by initial hand location (for example, see Fig. 5d), which is also coded in eye coordinates in this area²⁰. If both

target location and initial hand location are coded in eye coordinates in the PPC, then they can simply be subtracted to compute the target's location in hand coordinates, as shown in Fig. 1a, b. A role for the PPC in the comparison of target and hand position-related signals has been previously suggested^{21,22}, although the coordinate frame underlying this operation has not been specified until now. Simulations have shown that this operation can be implemented neurally by a simple, weighted summation of gain-modulated neurons like those found in the PRR (Fig. 5d)²³.

The fact that target locations are coded in both eye and hand coordinates in area 5 might mean that further processing is needed to complete the transformation to hand coordinates. However, it is also quite possible that this spatial representation is necessary for the functions performed by this cortical area, and therefore does not signify an intermediate step in the transformation process²⁴. Consistent with this latter possibility are the findings that area 5 receives visual, proprioceptive, and efference copy signals^{2–9}, which are probably represented in different coordinate frames²⁵. In principle, the gain-modulated cells in the PRR can be used to compute target locations purely in hand-centred coordinates, that is, without an eye-centred component, by weighting the convergence of activity differently. Such a scheme may be used to construct representations in other brain areas that are more heavily biased towards hand-centred coordinates than the one observed in area 5. □

Methods

Behavioural paradigm and neurophysiological recordings

Reaches were made to touch sensitive buttons that were 3.7 cm in diameter and set 7.5 cm apart, within a board placed 24 cm from the monkeys. In each condition, reaches were typically made to between 8 and 11 buttons located immediately surrounding the initial hand position and/or within an adjacent column of buttons (Fig. 2). Each button contained both a red and a green light-emitting diode (LED). The red LED instructed the animals where to direct and maintain their gaze, whereas the green LED instructed the animals where to place their hand. All trials began with the illumination of both a red and a green LED. A green (target) LED at another location was then briefly illuminated (300 ms duration). After a delay period of 600–1,000 ms, the LEDs instructing the initial hand location and fixation point were turned off and the animal reached to the remembered location of the target in complete darkness while maintaining fixation. Eye position was monitored using the scleral search coil technique. Reaching movements were made with the contralateral (left) arm.

We obtained recordings from the right hemispheres of two monkeys (*Macaca mulatta*). Eighty-nine area 5 neurons (61 from the animal CKY and 28 from animal DNT) were studied in all four conditions of the initial experiment. Furthermore, we studied 15 area 5 neurons and 17 PRR neurons (all from animal CKY) in the second experiment. The approximate centre of the PRR recording sites in this second experiment was 5-mm posterior to the centre of the area 5 sites, at depths below the superficial cortex. These sites represent only a portion of the larger reach-related region reported in an earlier study²⁶. All analyses were performed on the mean firing rate (5 repetitions) during a 400-ms epoch centred on movement onset.

Correlation analysis

Determination of statistical differences among correlations was facilitated by constructing a distribution of correlation coefficients ($N = 200$) for each coordinate frame(s), using standard statistical bootstrapping techniques²⁷. A Kruskal–Wallis test with nonparametric multiple comparisons ($P < 0.05$) was then performed on these distributions²⁸, which revealed that area 5 activity was best correlated when target locations were identical in both eye and hand coordinates (Fig. 3e).

At the population level, activity was generally well correlated in all reference frames, due in part to the fact that area 5 neurons exhibit relatively broad spatial tuning as well as the fact that cells with low firing rates tended to fire at low rates in all experimental conditions. To assess the degree to which our findings depended on the relative number of low firing pairs in Fig. 3a–e, we performed an additional analysis of the scatter in these plots, using a measure that is relatively insensitive to differing numbers of low-firing pairs and data points (which varied across Fig. 3a–e). More specifically, scatter (s) was quantified as

$$s = \sum_{i=1}^n \rho_i^2 / \sum_{i=1}^n r_i^2 \quad (1)$$

where ρ is the perpendicular distance to the unity line, and r is the distance to the origin. This analysis gave the same result as that which was obtained using the correlation coefficient, that is, scatter was least when target locations were identical in both eye and hand coordinates (Kruskal–Wallis test with nonparametric multiple comparisons, $P < 0.05$).

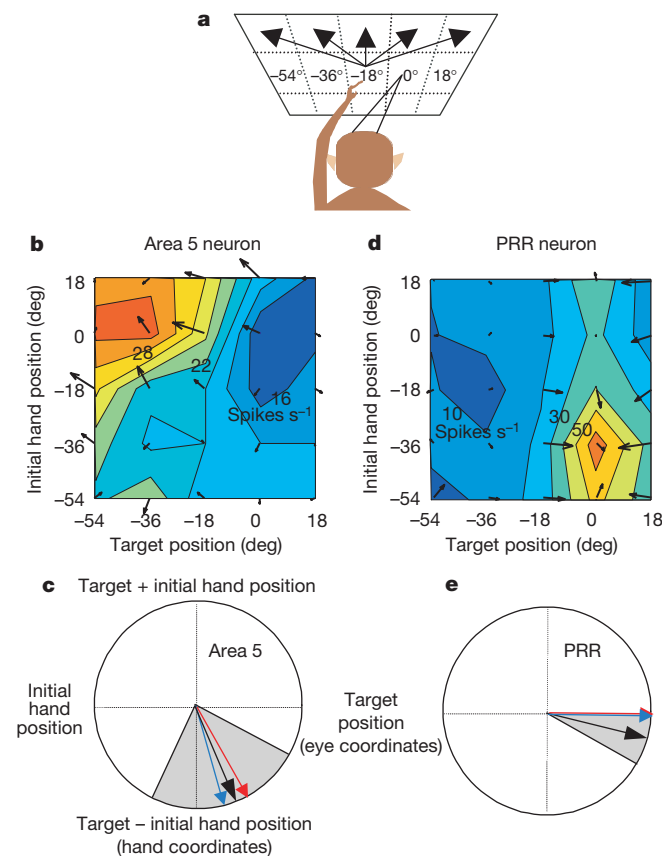


Figure 5 Results from the second experiment. **a**, Task schematic. **b**, Contour plot and gradient (black vectors) for one area 5 neuron. **c**, Resultant of the gradient in **b** (blue vector), population resultant of 15 single cell resultants (black vector), and gradient resultant for the idealized neuron in Fig. 4c (red vector). The shaded region indicates the 95% confidence interval for the population resultant²⁷. **d**, Contour plot and gradient for a PRR neuron. **e**, Resultant of the gradient in **d** (blue vector), resultant of 17 PRR single cell resultants (black vector), and gradient resultant for the idealized neuron in Fig. 4b (red vector).

Idealized neural responses

The mean firing rates (fr) of the idealized neurons shown in Fig. 4a–d, are given by equation (2) (from top to bottom, respectively):

$$\begin{aligned} fr &= e^{-\left(\frac{T^2}{2}\right)} \\ fr &= e^{-\left(\frac{T^2}{2} + \frac{H^2}{8}\right)} \\ fr &= e^{-\left(\frac{T^2}{4} + \frac{(T-H)^2}{4}\right)} \\ fr &= e^{-\frac{(T-H)^2}{2}} \end{aligned} \quad (2)$$

where T is the horizontal position of the target in eye coordinates, H is the horizontal position of the hand in eye coordinates, and $T - H$ is the horizontal position of the target in hand coordinates (see Fig. 1a). Our observations regarding these responses appear to be insensitive to both the form of these functions (gaussian versus sigmoid) as well as the nature of their interaction (multiplicative versus additive).

Gradient analysis

We estimated gradients from the data using an approximate numerical method (Matlab; Mathworks). The gradient resultant is a measure of the 'orientation' of an individual response field, and hence is an indicator of the variable or variables to which a neuron is most responsive (target position, initial hand position, and so on), regardless of the form of tuning (gaussian, sigmoid, and so on). To account for symmetrically shaped response fields we doubled the angles of the gradient vectors, then subtracted 360° from those angles greater than or equal to 360° , before taking the resultant²⁸. This procedure transformed the data in such a way that resultants could be expressed easily in terms of their dependence on target position and initial hand position, as well as their sum and difference (Fig. 5c, e). Resultants could not, however, be mapped directly onto the response fields from which they were derived. For example, although neurons coding target position purely in hand-centred coordinates have obliquely oriented response fields (Fig. 4d), as points along the unity line correspond to identical hand-centred target positions, their gradient resultants would be expected to point straight down in Fig. 5c, e. Single cell and population resultants were normalized to unit length before plotting in Fig. 5.

Received 3 September 2001; accepted 25 January 2002.

- Rondot, P., Recondo, J. & de Ribadeau Dumas, J. Visuomotor ataxia. *Brain* **100**, 355–376 (1977).
- Kalaska, J. F., Caminiti, R. & Georgopoulos, A. P. Cortical mechanisms related to the direction of two-dimensional arm movements—relations in parietal area 5 and comparison with motor cortex. *Exp. Brain Res.* **51**, 247–260 (1983).
- Georgopoulos, A. P., Caminiti, R. & Kalaska, J. F. Static spatial effects in motor cortex and area 5: quantitative relations in a two-dimensional space. *Exp. Brain Res.* **4**, 446–454 (1984).
- Kalaska, J. F., Cohen, D. A. D., Prud'homme, M. & Hyde, M. L. Parietal area 5 neuronal activity encodes movement kinematics, not movement dynamics. *Exp. Brain Res.* **80**, 351 (1990).
- Lacquaniti, F., Guignon, E., Bianchi, L., Ferraina, S. & Caminiti, R. Representing spatial information for limb movement: role of area 5 in the monkey. *Cerebr. Cortex* **5**, 391–409 (1995).
- Kalaska, J. F. & Crammond, D. J. Deciding not to go: neuronal correlates of response selection in a go/nogo task in primate premotor and parietal cortex. *Cerebr. Cortex* **5**, 410–428 (1995).
- Scott, S. H., Sergio, L. H. & Kalaska, J. F. Reaching movements with similar hand paths but different arm orientations. II. Activity of individual cells in dorsal premotor cortex and parietal area 5. *J. Neurophysiol.* **78**, 2413–2426 (1997).
- Graziano, M. S. A., Cooke, D. F. & Taylor, C. S. R. Coding the location of the arm by sight. *Science* **290**, 1782–1786 (2000).
- Battaglia-Mayer, A. *et al.* Early coding of reaching in the parietooccipital cortex. *J. Neurophysiol.* **83**, 2374–2391 (2000).
- Flanders, M., Tillery, S. I. H. & Soechting, J. F. Early stages in a sensorimotor transformation. *Behav. Brain Sci.* **15**, 309–320 (1992).
- Henriques, D. Y. P., Klier, E. M., Smith, M. A., Lowy, D. & Crawford, J. D. Gaze-centered remapping of remembered visual space in an open-loop pointing task. *J. Neurosci.* **18**, 1583–1594 (1998).
- McIntyre, J., Stratta, F. & Lacquaniti, F. Short-term memory for reaching to visual targets: psychophysical evidence for body-centered reference frames. *J. Neurosci.* **18**, 8423–8435 (1998).
- Jordan, M. I. & Rumelhart, D. E. Forward models: supervised learning with a distal teacher. *Cognitive Sci.* **16**, 307–354 (1992).
- Batista, A. P., Buneo, C. A., Snyder, L. H. & Andersen, R. A. Reach plans in eye-centered coordinates. *Science* **285**, 257–260 (1999).
- Strick, P. L. & Kim, C. C. Input to primate motor cortex from posterior parietal cortex (area 5). I. Demonstration by retrograde transport. *Brain Res.* **157**, 325–330 (1978).
- Wiesendanger, R., Wiesendanger, M. & Ruegg, D. G. An anatomical investigation of the corticopontine projection in the primate (*Macaca fascicularis* and *Saimiri sciureus*). II. The projection from the frontal and parietal areas. *Neuroscience* **4**, 747–765 (1979).
- Petrides, M. & Pandya, D. N. Projections to the frontal cortex from the posterior parietal region in the rhesus monkey. *J. Comp. Neurol.* **228**, 105–116 (1984).
- Carozza, M., McIntyre, J., Zago, M. & Lacquaniti, F. Viewer-centered and body-centered frames of reference in direct visuomotor transformations. *Exp. Brain Res.* **129**, 201–210 (1999).
- Andersen, R. A., Essick, G. K. & Siegel, R. M. Encoding of spatial location by posterior parietal neurons. *Science* **230**, 456–458 (1985).
- Buneo, C. A., Batista, A. P. & Andersen, R. A. Frames of reference for reach-related activity in two parietal areas. *Soc. Neurosci. Abstr.* **24**, 262 (1998).
- Desmurget, M. *et al.* Role of the posterior parietal cortex in updating reaching movements to a visual target. *Nature Neurosci.* **2**, 563–567 (1999).
- Bullock, D., Cisek, P. & Grossberg, S. Cortical networks for control of voluntary arm movements under variable force conditions. *Cerebr. Cortex* **8**, 48–62 (1998).
- Salinas, E. & Abbott, L. Transfer of coded information from sensory to motor networks. *J. Neurosci.*

- 15, 6461–6474 (1995).
24. Deneve, S., Latham, P. E. & Pouget, A. Efficient computation and cue integration with noisy population codes. *Nature Neurosci.* **4**, 826–831 (2001).
25. Moran, D. W. & Schwartz, A. B. Motor cortical representation of speed and direction during reaching. *J. Neurophysiol.* **82**, 2676–2692 (1999).
26. Snyder, L. H., Batista, A. P. & Andersen, R. A. Coding of intention in the posterior parietal cortex. *Nature* **386**, 167–170 (1997).
27. Efron, B. & Tibshirani, R. J. *An Introduction to the Bootstrap* (Chapman & Hall, New York, 1993).
28. Zar, J. H. *Biostatistical Analysis* (Prentice-Hall, Upper Saddle River, New Jersey, 1999).
29. Andersen, R. A., Snyder, L. H., Li, C. S. & Stricanne, B. Coordinate transformations in the representation of spatial information. *Curr. Opin. Neurobiol.* **3**, 171–176 (1993).

Acknowledgements

This work was supported by the Defense Advanced Research Projects Agency (DARPA), the National Eye Institute, the Sloan-Schwartz Center for Theoretical Neurobiology, the James G. Boswell Foundation and an NIH training grant fellowship to C.A.B. We thank B. Gillikin and V. Shcherbatyuk for technical assistance; D. Dubowitz for collecting and processing the MRI data; J. Baer and J. Wynne for veterinary care; and C. Reyes-Marks for administrative assistance. We also thank J. Boline and K. Shenoy for comments.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to R.A.A. (e-mail: andersen@vis.caltech.edu).

Chondroitinase ABC promotes functional recovery after spinal cord injury

Elizabeth J. Bradbury*, Lawrence D. F. Moon†§, Reena J. Popat*, Von R. King‡, Gavin S. Bennett*, Preena N. Patel*, James W. Fawcett† & Stephen B. McMahon*

* Sensory Function Group, Centre for Neuroscience Research, Hodgkin Building, Kings College London, Guy's Campus, London Bridge, London SE1 1UL, UK

† Department of Physiology and Centre for Brain Repair, University of Cambridge, Robinson Way, Cambridge CB2 2PY, UK

‡ Department of Neuroscience, St Bartholomew's and the Royal London School of Medicine and Dentistry, Queen Mary, University of London, Mile End Road, London E1 4NS, UK

The inability of axons to regenerate after a spinal cord injury in the adult mammalian central nervous system (CNS) can lead to permanent paralysis. At sites of CNS injury, a glial scar develops, containing extracellular matrix molecules including chondroitin sulphate proteoglycans (CSPGs)^{1,2}. CSPGs are inhibitory to axon growth *in vitro*^{3–5}, and regenerating axons stop at CSPG-rich regions *in vivo*⁶. Removing CSPG glycosaminoglycan (GAG) chains attenuates CSPG inhibitory activity^{7–10}. To test the functional effects of degrading chondroitin sulphate (CS)-GAG after spinal cord injury, we delivered chondroitinase ABC (ChABC) to the lesioned dorsal columns of adult rats. We show that intra-thecal treatment with ChABC degraded CS-GAG at the injury site, upregulated a regeneration-associated protein in injured neurons, and promoted regeneration of both ascending sensory projections and descending corticospinal tract axons. ChABC treatment also restored post-synaptic activity below the lesion after electrical stimulation of corticospinal neurons, and promoted functional recovery of locomotor and proprioceptive behaviours. Our results demonstrate that CSPGs are important

§ Present address: Miami Project to Cure Paralysis, 1095 NW 14th Terrace, PO Box 16960, Mail Locator R-48, Miami, Florida 33101, USA.

inhibitory molecules *in vivo* and suggest that their manipulation will be useful for treatment of human spinal injuries.

Adult rats received a cervical (C4) dorsal column crush lesion and bolus intrathecal infusions of protease-free ChABC or control solution. We assessed the degradation of CS-GAG at the injury site with the antibody 2B6, which recognizes an epitope on CSPG core proteins but not intact CS-GAG¹⁰. In control animals (unlesioned or lesioned and treated with vehicle) no 2B6 immunoreactivity was apparent (Fig. 1a). However, *in vitro* ChABC incubation of adjacent lesioned sections revealed extensive immunoreactivity (Fig. 1b), confirming the presence of CSPGs around the injury site. Lesioned rats treated *in vivo* with ChABC showed intense 2B6 immunoreactivity around the lesion site and in white matter tracts extending at least 4 mm rostrally and caudally (Fig. 1c, d). Thus, *in vivo* delivery of ChABC successfully removed GAG chains from CSPGs at and around the injury site.

Growth-associated protein 43 (GAP-43) is upregulated in primary sensory neurons after peripheral, but not central, nerve injury^{11,12}, and thus is associated with a regenerative state. We assessed GAP-43 expression after ChABC treatment in C5/6 dorsal root ganglia (DRG). In sham-operated controls, GAP-43 was present in 25% of small-diameter (less than 40 μm) and 6% of large-diameter (greater than 40 μm) neurons (Fig. 2a), and

expression was similar in lesioned animals treated with vehicle (Fig. 2b). However, ChABC treatment increased GAP-43 in large-diameter cells (which project through the lesion site) to 22% ($P < 0.001$, analysis of variance (ANOVA), Fig. 2c). Thus, degrading GAG components of CSPGs promotes GAP-43 expression in injured neurons, indicating increased regenerative propensity in these cells.

Anatomical regeneration of injured dorsal column axons was assessed by labelling forelimb nerves with cholera toxin B-subunit (CTB). After lesion and infusion of vehicle, retraction of labelled fibres occurred, with few approaching, and none entering the lesion site (Fig. 2d). In contrast, in all lesioned animals that received ChABC infusions, thick fibre bundles approached the lesion (Fig. 2e). Notably, many axons traversed lesioned tissue, either

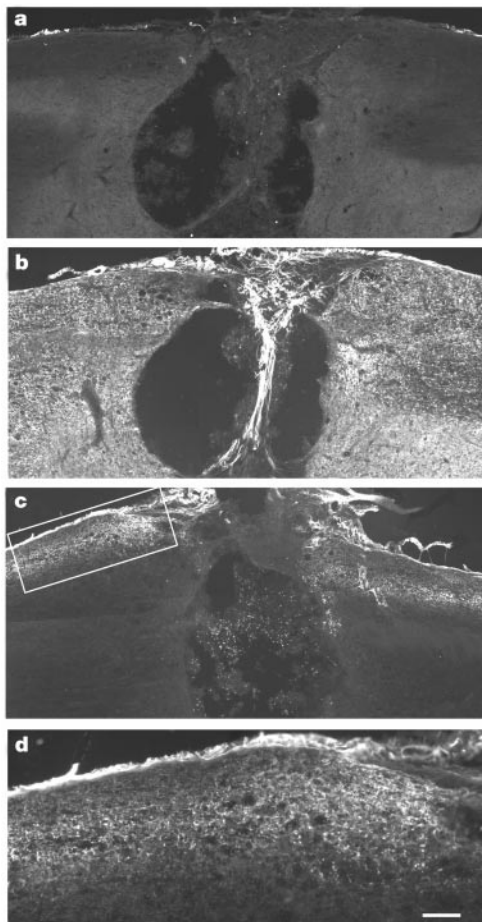


Figure 1 ChABC degrades CS-GAG *in vivo*. **a**, No 2B6 immunoreactivity is apparent in spinal cord sections after lesion and control infusions. **b**, Adjacent tissue incubated *in vitro* with ChABC shows intense immunoreactivity for 2B6 throughout the entire section, indicating an abundance of CSPGs in and around a site of spinal cord injury. **c**, After *in vivo* infusions of ChABC, 2B6 immunoreactivity is apparent at the lesion site and in the white matter, where spinal cord pathways project. **d**, High power of boxed area in **c**. Scale bar, 200 μm (**a-c**); 80 μm (**d**).

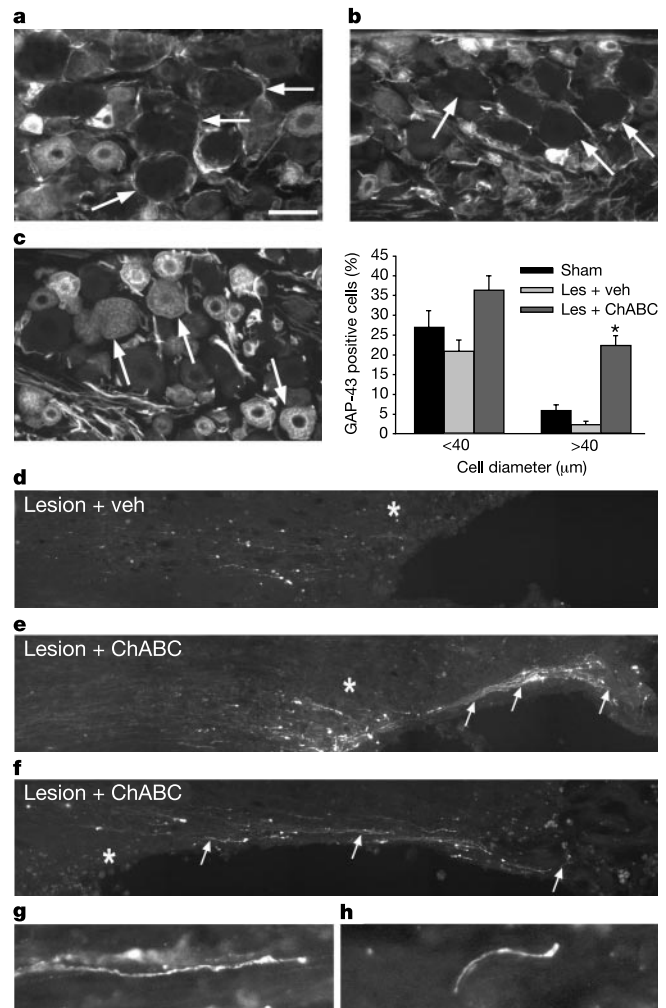


Figure 2 ChABC induces upregulation of GAP-43 in lesioned DRG neurons and regeneration of dorsal column axons. **a, b**, In control rats, GAP-43 protein expression is seen in a limited population of cervical DRG neurons (**a**), and does not change after lesion and control infusions (**b**; arrows in **a, b** indicate large-diameter neurons that do not express GAP-43). **c**, In contrast, a marked upregulation of GAP-43 expression is apparent after ChABC treatment, particularly in large-diameter DRG neurons (arrows). Quantification data (right) are means $\pm$ s.e.m. (asterisk denotes a significant difference from controls, $P < 0.001$, one-way ANOVA, Tukey post-hoc). Les, lesion; veh, vehicle. **d**, Axon tracing in the spinal cord reveals few CTB-labelled fibres proximal to the lesion after control infusions, with retraction bulbs indicating abortive regeneration. **e, f**, In contrast, after ChABC treatment many axons approach the lesion site and bundles of regenerating fibres traverse the lesioned tissue (arrows). **g, h**, Growth-cone-like endings are observed on regenerating fibres in lesioned animals treated with ChABC. Asterisks in **d-f**, indicate the lesion site. Scale bar, 50 μm (**a-c**); 400 μm (**d-f**); 20 μm (**g, h**).

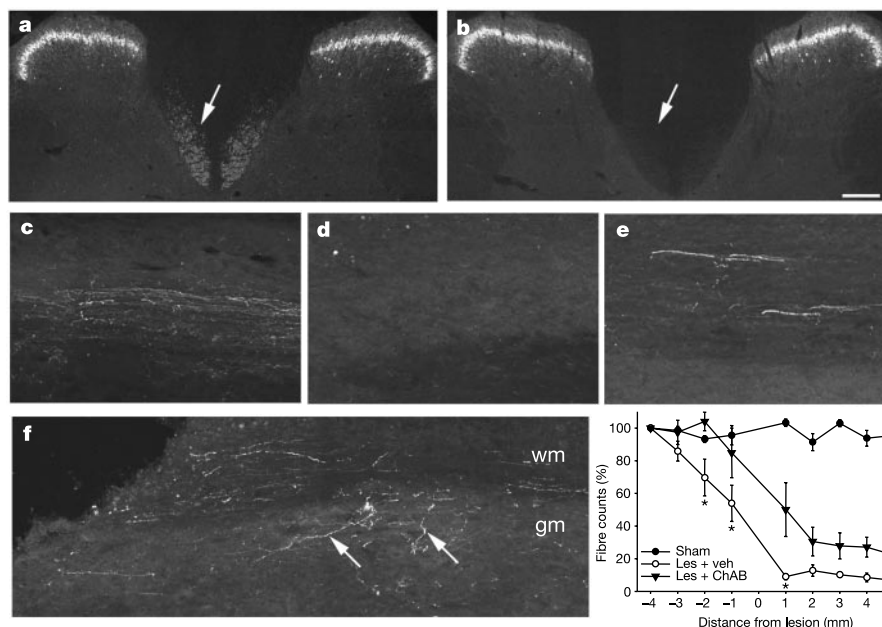


Figure 3 ChABC promotes regeneration of corticospinal tract axons. **a**, In control lumbar spinal cord, PKC- γ immunoreactivity is present in the dorsal horn and the CST. **b**, After a dorsal column lesion, PKC- γ immunoreactivity is no longer present in the CST (arrows in **a** and **b**), confirming a complete lesion. **c**, BDA-labelled CST fibres in the cervical spinal cord of controls. **d**, **e**, In lesioned animals BDA-labelled fibres are not observed below the injury site after vehicle infusions (**d**), but are present after ChABC infusions (**e**). **f**, In a

ChABC-treated animal many axons are observed at the lesion site and appear to send collaterals (arrows) from white matter (wm) to grey matter (gm), indicating terminal arborization of regenerating axons. Quantification data (right) are percentages (means $\pm$ s.e.m.; asterisks denote significant difference between vehicle and ChABC treatment, $P < 0.05$, two-way ANOVA, Tukey post-hoc). Scale bar, 200 μ m (**a–e**); 100 μ m (**f**). Les, lesion; veh, vehicle.

growing around cavities or within the lesion epicentre (Fig. 2e, f). Growth-cone-like endings were apparent on axon tips in rats infused with ChABC (Fig. 2g, h), indicative of regeneration. Bundles of axons grew up to 2 mm through lesioned tissue, and many individual fibres were apparent 4 mm past the lesion site, indicating robust regeneration of dorsal column axons.

Regeneration of the major descending pathway in the dorsal columns, the corticospinal tract (CST), was also investigated. In unlesioned lumbar spinal cord, protein kinase C (PKC)- γ immunoreactivity (a marker of the CST¹³) was observed in the dorsal columns and lamina II (Fig. 3a). In all lesioned animals, no PKC- γ immunoreactivity was observed in the CST at lumbar levels (Fig. 3b), confirming complete transection of this tract. Regeneration of the CST was assessed using biotinylated dextran amine (BDA) injected into the motor cortex. In unlesioned animals, a thick fibre bundle was observed in cervical spinal cord (Fig. 3c). After dorsal column injury, CST axons had retracted from the lesion¹⁴, and no fibres were seen beyond the lesion (Fig. 3d). However, ChABC treatment prevented CST retraction and promoted regeneration, with significantly more fibres seen at and below the lesion compared with vehicle treatment ($P < 0.001$, ANOVA, Fig. 3e). In ChABC-treated animals, some axons sent arborizing collaterals into grey matter (Fig. 3f).

Electrophysiological experiments determined that regenerated CST axons established functional connections. In anaesthetized control animals, electrical stimuli applied to motor cortex evoked large postsynaptic cord dorsum potentials (CDPs), as described previously¹⁵, with an average latency of 3.9 ± 0.4 ms at C4 (Fig. 4). Acute dorsal column lesions largely abolished these CDPs below the lesion. The small remaining response presumably represents activation via non-dorsal column motor pathways. In vehicle-treated animals, CDPs were present above, but only minimally below the lesion ($P < 0.05$, ANOVA, Fig. 4). In contrast, ChABC-treated animals showed large CDPs up to 7 mm below the lesion, averaging about 40% of C3 CDP magnitude, significantly different from responses in vehicle-treated preparations ($P < 0.001$, ANOVA,

Fig. 4). Responses had normal configurations, but were delayed (mean latency of 18.2 ± 2.6 ms, $P < 0.005$ compared with controls, t -test, Fig. 4), consistent with poor re-myelination of regenerating axons¹⁶. All ChABC-treated animals showed similar responses, which were abolished by acute dorsal column re-section; strong evidence of new connections of regenerated CST axons.

Dorsal column projections are important for discriminative touch and proprioception, and—together with local spinal reflexes—for skilled motor function. Rats were assessed by several behavioural tasks of forelimb function requiring sensorimotor skills. Adhesive tape-removal tasks assessed sensory (awareness of the tape) and motor (ability to remove the tape) function (Fig. 5a, b); unlesioned sham controls could perform these tasks quickly. Lesioned rats were severely impaired with or without ChABC treatment. These impairments persisted throughout six weeks of testing, consistent with the failure of ChABC treatment to promote regeneration to hindbrain sensory nuclei (6 weeks post-lesion latencies were: 7.4 ± 1.4 s and 3.8 ± 0.5 s (sham); 54.9 ± 3.1 s and 48.0 ± 4.4 s (lesion plus vehicle); 41.7 ± 8.3 s and 40.5 ± 11.1 s (lesion plus ChABC), for sense and removal tasks, respectively, $P < 0.001$, ANOVA, Fig. 5a, b). The number of forelimb foot slips were recorded when rats crossed a narrow beam or grid. In both tasks, unlesioned sham controls made few foot slips but vehicle-treated lesioned rats were severely and significantly impaired throughout testing, albeit with some recovery over time¹⁷ ($P < 0.001$, ANOVA; at 6 weeks after lesion, foot slips had increased compared with sham controls, from 0 ± 0 to 21.1 ± 4.2 and 0.8 ± 0.5 to 8.8 ± 1.9 on the beam and grid, respectively, Fig. 5c, d). ChABC treatment produced a marked recovery of function, with improvements from 2 weeks (beam) and 1 week (grid) onwards (not significantly different from sham controls, $P > 0.1$, ANOVA; at 6 weeks after lesion, foot slips were 2.6 ± 0.8 and 1.6 ± 0.7 on the beam and grid, respectively, Fig. 5c, d). Walking patterns were then assessed by analysing footprint spacing. Vehicle-treated lesioned rats took significantly shorter and wider strides than sham controls (stride length, 118.1 ± 6.3 mm versus

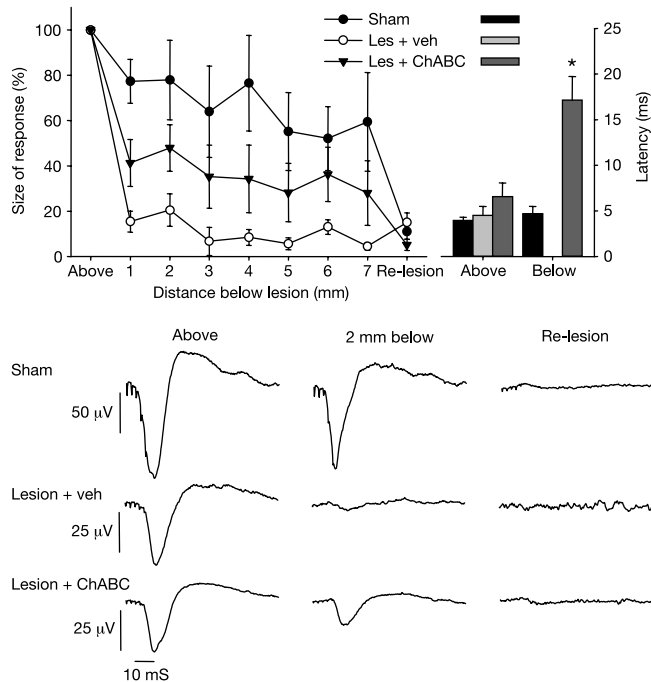


Figure 4 ChABC restores functional connections of descending motor pathways. Graph shows the average size of cortical evoked cord dorsum potentials (CDPs) 1 segment above C4 (lesion site) and at 1 mm intervals caudal to this site (data normalized to the size of the rostral recording). ChABC treatment increases CDPs below the lesion ($P < 0.001$, two-way ANOVA, compared with vehicle). Representative sample traces show that cortical stimulation normally evokes a short-latency negative-positive wave, which is absent 2 mm below an acute or chronic dorsal column lesion, but is partially restored by ChABC treatment. Notably, this restored activity is abolished by a re-lesion, indicating CST regeneration. The bar graph shows no differences in response latencies above the lesion ($P > 0.1$, one-way ANOVA) but below the lesion (measurement not possible in vehicle-treated preparations due to absence of evoked activity) latencies are longer in ChABC-treated animals (asterisk denotes significant difference from sham group, $P < 0.005$, *t*-test).

148.9 $\pm$ 4.8 mm; width, 25.2 $\pm$ 3.7 mm versus 12.5 $\pm$ 1.9 mm; 6 weeks after lesion, $P < 0.02$, ANOVA, Fig. 5e, f). ChABC treatment largely prevented changes (length and width were 137.1 $\pm$ 11.0 mm and 15.4 $\pm$ 3.6 mm, respectively, not significantly different from shams, $P > 0.1$, ANOVA, Fig. 5e, f).

Our results demonstrate, using anatomical, electrophysiological and behavioural outcomes, that ChABC promotes regeneration and restores function after spinal cord injury. Several CSPGs, including neurocan, phosphacan and NG2, are upregulated at CNS injury sites^{18–21}, and we have shown that CS-GAG chains contribute towards regenerative failure. Anatomical regeneration was limited, presumably due to other inhibitory factors in lesioned spinal cord^{22,23}. However, we observed very clear recovery of function after ChABC treatment, for which there are several possible anatomical substrates. First, a limited regrowth of sensory and CST motor axons might account, by means of new local segmental connections, for the behavioural recovery. Second, CSPGs may inhibit intact pathways from sprouting, allowing ChABC treatment to result in beneficial compensatory sprouting mechanisms, as reported for IN-1 treatment²⁴. Chondroitinase ABC, and other potential treatments that affect CSPG production after injury, join blockade of NogoA, neurotrophic factor treatment, gene therapy and cellular grafting^{25–28} as interventions that promote spinal cord regeneration, and may have therapeutic potential for the treatment of human spinal cord injuries. □

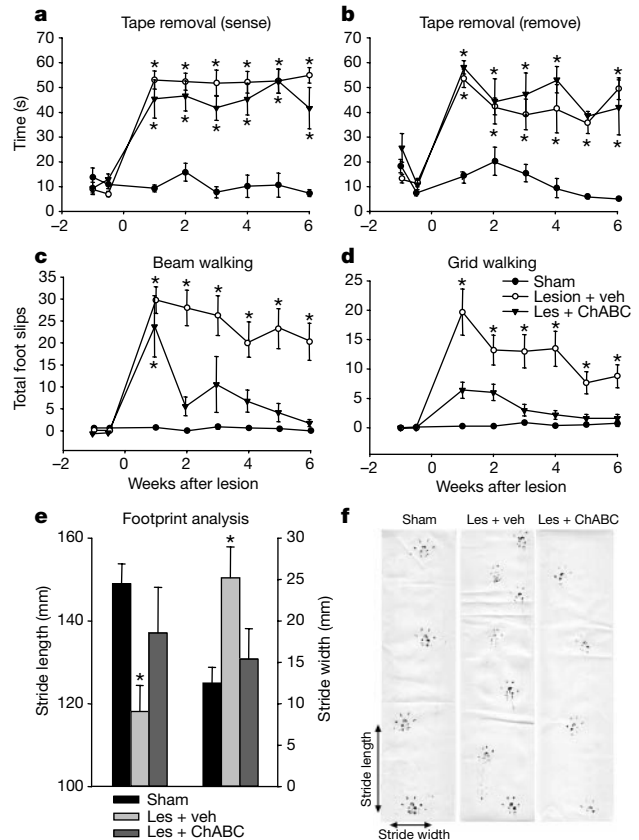


Figure 5 Functional recovery after ChABC treatment. **a, b**, In a tape removal task, lesioned rats treated with either vehicle solution or ChABC were severely impaired, compared with sham controls, both in their ability to sense and subsequently remove the adhesive tape. **c, d**, In the beam and grid tasks, lesioned rats treated with vehicle made significantly more foot-slip errors than controls. In contrast, lesioned rats treated with ChABC made a marked functional recovery on both these tasks. **e**, Footprint analysis revealed impaired walking patterns in lesioned rats treated with vehicle, with the strides becoming shorter and wider than controls; again, functional recovery was observed in lesioned rats treated with ChABC, with both the length and width of strides taken not significantly different from controls. **f**, Representative footprints from each group. Data are mean $\pm$ s.e.m. (asterisks denote significant difference from sham controls, $P < 0.02$, ANOVA, Tukey post-hoc).

Methods

Spinal cord injury and ChABC treatment

Anaesthetized (sodium pentobarbital, 40 mg kg⁻¹) adult male Wistar rats received a bilateral lesion to the dorsal columns²⁹ at spinal level C4. Concurrently, a silastic tube was inserted intrathecally to lie just rostral to the lesion site, and externalized to deliver bolus injections of high-purity, protease-free chondroitinase ABC (ChABC; Seikagaku Corporation). Immediately after the lesion, 6 μ l ChABC (10 U ml⁻¹) was injected followed by a 6- μ l saline flush ($n = 17$). A further group received the spinal cord lesion with either saline or control enzyme treatment (Penicillinase, Sigma; same μ g protein delivered, $n = 21$). ChABC or control solution was delivered on alternate days for 10 days after lesion. Control animals received sham surgery ($n = 20$). Rats were perfused at 2 weeks after lesion to check effectiveness of treatment in removing CS-GAG ($n = 4$ per group), tested behaviourally over 6 weeks ($n = 5, 9$ and 8 for ChABC, vehicle and sham, respectively), then perfused at 8 weeks after lesion for anatomical analysis of ascending systems ($n = 4$ per group), perfused at 10 weeks after lesion for analysis of CST regeneration ($n = 4$ per group), or used for terminal electrophysiology experiments at 13–17 days after lesion ($n = 4$ per group).

Confirmation of CS-GAG digestion

Tissue processing was as described previously²⁹. Parasagittal sections (20 μ m) of cervical spinal cord were immunostained using monoclonal antibody 2B6 (Seikagaku Corporation, 1:5,000), with tyramide signal amplification (NEN). Some tissue sections from lesioned cords treated with vehicle *in vivo* were incubated *in vitro* with ChABC (0.2 U ml⁻¹, 2 h, 37 °C). We processed tissue from all experimental conditions in parallel.

GAP-43 analysis in DRG neurons

Sections of C5 and C6 DRG (10 μ m) were double-immunostained for β III tubulin (Promega, 1:1000) to identify all neurons, and for the growth-associated protein GAP-43 (gift from G. Wilkin, 1:2000), using 7-amino-4-methylcoumarin-3-acetic-acid- and tetramethylrhodamine isothiocyanate (TRITC)-conjugated secondary antibodies. The percentage of GAP-43-positive cells was determined in four sections per animal.

Axon tract tracing in the spinal cord

Ascending dorsal column axons were labelled using the CTB tracer²⁹, injected into the left median nerve of anaesthetized rats. Longitudinal sections (20 μ m) were immunostained for glial fibrillary acidic protein, to identify the lesion site, and for CTB, to identify the labelled dorsal column axons²⁹. Descending CST axons were labelled with 1 μ l BDA (Molecular Probes; 10% in saline) injected under anaesthesia at four sites over the left primary motor cortex, 1 mm below the dorsal surface of the brain. BDA-labelled fibres were visualized in parasagittal spinal cord sections (20 μ m) with extra-avidin conjugated to fluorescein isothiocyanate. All BDA-labelled fibres observed within a 1-mm square grid were counted at measured intervals from 4 mm above to 5 mm below the lesion site by an experimenter, blinded to treatment. Due to variability in labelling, axon numbers were calculated as a percentage of the fibres seen 4 mm above the lesion, where the CST was intact. To confirm a complete CST lesion, transverse lumbar spinal cord sections (20 μ m) were immunostained with an antibody against the γ -subunit of protein kinase C (PKC- γ , Santa Cruz, 1:1000), visualized with a TRITC-conjugated secondary antibody.

Electrophysiology

In terminal electrophysiological experiments, the sensory motor cortex and cervical spinal cord were exposed in urethane-anaesthetized (1.5 g kg⁻¹) rats. Cortical evoked potentials were elicited by electrical stimulation (five square-wave pulses at 400 Hz, 100 μ A, 200 μ s, delivered every 2 s) of the left sensory motor cortex using a 0.5-mm concentric needle electrode lowered 1 mm into the cortex. For each experiment we mapped the optimal stimulation site, located 1–2 mm lateral and 1 mm rostrocaudal from Bregma. Postsynaptic potentials evoked by the cortical stimuli were recorded with a silver ball electrode placed medially on the contralateral cord surface. At each recording site (1 segment above and 1–7 mm below the lesion at C4), 64 responses were averaged and stored for off-line analysis of response magnitude (area under curve) and latency.

Behavioural assessment

After baseline testing, we tested animals once a week for 6 weeks after lesion. We averaged right and left forepaw scores, as no differences were observed between them. Experimenters were blind to the treatment. A tape removal test (adapted from ref. 24) produced separate scores for sensory and motor behaviour. Adhesive tape (0.3 inch $\times$ 1 inch) was placed on the forepaw, and the time taken to sense the presence of the tape (indicated by paw shake) was determined. For animals that sensed the tape, the removal time was also scored. Rats were tested on two locomotor tasks requiring sensorimotor integration (adapted from ref. 30). Rats were trained to cross a narrow metal beam (1.25 inch $\times$ 36 inches) and a wire grid (12 inches $\times$ 36 inches with 1 inch $\times$ 1 inch grid squares). Forepaw foot slips were recorded (determined by a paw slipping off the beam or below the plane of the grid). In a footprint analysis (adapted from ref. 30), rat forepaws were covered with ink to record walking patterns during continuous locomotion across a wooden runway (4 inches $\times$ 36 inches), and stride length and width were calculated.

Received 4 September 2001; accepted 8 February 2002.

1. Fawcett, J. W. & Asher, R. A. The glial scar and central nervous system repair. *Brain Res. Bull.* **49**, 377–391 (1999).
2. Fitch, M. T. & Silver, J. *CNS Regeneration: Basic Science and Clinical Advances* (eds Tuszynski, M. H. & Kordower, J. H.) 55–88 (Academic, San Diego, 1999).
3. McKeon, R. J., Schreiber, R. C., Rudge, J. S. & Silver, J. Reduction of neurite outgrowth in a model of glial scarring following CNS injury is correlated with the expression of inhibitory molecules on reactive astrocytes. *J. Neurosci.* **11**, 3398–3411 (1991).
4. Smith-Thomas, L. C. *et al.* An inhibitor of neurite outgrowth produced by astrocytes. *J. Cell Sci.* **107**, 1687–1695 (1994).
5. Niederost, B. P., Zimmermann, D. R., Schwab, M. E. & Bandtlow, C. E. Bovine CNS myelin contains neurite growth-inhibitory activity associated with chondroitin sulfate proteoglycans. *J. Neurosci.* **19**, 8979–8989 (1999).
6. Davies, S. J., Goucher, D. R., Doller, C. & Silver, J. Robust regeneration of adult sensory axons in degenerating white matter of the adult rat spinal cord. *J. Neurosci.* **19**, 5810–5822 (1999).
7. McKeon, R. J., Hoke, A. & Silver, J. Injury-induced proteoglycans inhibit the potential for laminin-mediated axon growth on astrocytic scars. *Exp. Neurol.* **136**, 32–43 (1995).
8. Zuo, J., Neubauer, D., Dyess, K., Ferguson, T. A. & Muir, D. Degradation of chondroitin sulfate proteoglycan enhances the neurite-promoting potential of spinal cord tissue. *Exp. Neurol.* **154**, 654–662 (1998).
9. Fidler, P. S. *et al.* Comparing astrocytic cell lines that are inhibitory or permissive for axon growth: the major axon-inhibitory proteoglycan is NG2. *J. Neurosci.* **19**, 8778–8788 (1999).
10. Moon, L. D., Asher, R. A., Rhodes, K. E. & Fawcett, J. W. Regeneration of CNS axons back to their target following treatment of adult rat brain with chondroitinase ABC. *Nature Neurosci.* **4**, 465–466 (2001).
11. Chong, M. S. *et al.* GAP-43 expression in primary sensory neurons following central axotomy. *J. Neurosci.* **14**, 4375–4384 (1994).
12. Bradbury, E. J., McMahon, S. B. & Ramer, M. S. Keeping in touch: sensory neurone regeneration in the CNS. *Trends Pharmacol. Sci.* **21**, 389–394 (2000).
13. Mori, M., Kose, A., Tsujino, T. & Tanaka, C. Immunocytochemical localization of protein kinase C subtypes in the rat spinal cord: light and electron microscopic study. *J. Comp Neurol.* **299**, 167–177 (1990).
14. Hill, C. E., Beattie, M. S. & Bresnahan, J. C. Degeneration and sprouting of identified descending

- supraspinal axons after contusive spinal cord injury in the rat. *Exp. Neurol.* **171**, 153–169 (2001).
15. Wall, P. D. & Lidieth, M. Five sources of a dorsal root potential: their interactions and origins in the superficial dorsal horn. *J. Neurophysiol.* **78**, 860–871 (1997).
16. Ramer, M. S., Priestley, J. V. & McMahon, S. B. Functional regeneration of sensory axons into the adult spinal cord. *Nature* **403**, 312–316 (2000).
17. Murray, M. Strategies and mechanisms of recovery after spinal cord injury. *Adv. Neurol.* **72**, 219–225 (1997).
18. Levine, J. M. Increased expression of the NG2 chondroitin-sulfate proteoglycan after brain injury. *J. Neurosci.* **14**, 4716–4730 (1994).
19. McKeon, R. J., Juryne, M. J. & Buck, C. R. The chondroitin sulfate proteoglycans neurocan and phosphacan are expressed by reactive astrocytes in the chronic CNS glial scar. *J. Neurosci.* **19**, 10778–10788 (1999).
20. Asher, R. A. *et al.* Neurocan is upregulated in injured brain and in cytokine-treated astrocytes. *J. Neurosci.* **20**, 2427–2438 (2000).
21. Plant, G. W., Bates, M. L. & Bunge, M. B. Inhibitory proteoglycan immunoreactivity is higher at the caudal than the rostral Schwann cell graft-transected spinal cord interface. *Mol. Cell Neurosci.* **17**, 471–487 (2001).
22. Bregman, B. S. *et al.* Recovery from spinal cord injury mediated by antibodies to neurite growth inhibitors. *Nature* **378**, 498–501 (1995).
23. Pasterkamp, R. J., Anderson, P. N. & Verhaagen, J. Peripheral nerve injury fails to induce growth of lesioned ascending dorsal column axons into spinal cord scar tissue expressing the axon repellent Semaphorin3A. *Eur. J. Neurosci.* **13**, 457–471 (2001).
24. Thallmair, M. *et al.* Neurite growth inhibitors restrict plasticity and functional recovery following corticospinal tract lesions. *Nature Neurosci.* **1**, 124–131 (1998).
25. Ramer, M. S., Harper, G. P. & Bradbury, E. J. Progress in spinal cord research—a refined strategy for the International Spinal Research Trust. *Spinal Cord* **38**, 449–472 (2000).
26. Fouad, K., Dietz, V. & Schwab, M. E. Improving axonal growth and functional recovery after experimental spinal cord injury by neutralizing myelin associated inhibitors. *Brain Res. Brain Res. Rev.* **36**, 204–212 (2001).
27. Jones, L. L., Oudega, M., Bunge, M. B. & Tuszynski, M. H. Neurotrophic factors, cellular bridges and gene therapy for spinal cord injury. *J. Physiol.* **533**, 83–89 (2001).
28. Murray, M. & Fischer, I. Transplantation and gene therapy: combined approaches for repair of spinal cord injury. *Neuroscientist* **7**, 28–41 (2001).
29. Bradbury, E. J. *et al.* NT-3 promotes growth of lesioned adult rat sensory axons ascending in the dorsal columns of the spinal cord. *Eur. J. Neurosci.* **11**, 3873–3883 (1999).
30. Kunkel-Bagden, E., Dai, H. N. & Bregman, B. S. Methods to assess the development and recovery of locomotor function after spinal cord injury in rats. *Exp. Neurol.* **119**, 153–164 (1993).

Acknowledgements

We thank T. J. Boucher for critical review of the manuscript and J. V. Priestley, M. Lidieth, B. J. Kerr and M. S. Ramer for their help. This work was supported by The Wellcome Trust, MRC, International Spinal Research Trust and Action Research.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to E.J.B. (e-mail: elizabeth.bradbury@kcl.ac.uk).

The *Drosophila* immune response against Gram-negative bacteria is mediated by a peptidoglycan recognition protein

Marie Gottar*, Vanessa Gobert*, Tatiana Michel*, Marcia Belvin†, Geoffrey Duyk†, Jules A. Hoffmann*, Dominique Ferrandon*‡ & Julien Royet*‡

* UPR 9022 du Centre National de la Recherche Scientifique, Institut de Biologie Moléculaire et Cellulaire, 15 rue René Descartes, F67084 Strasbourg, Cedex, France

† Exelixis Inc., South San Francisco, California 94083, USA

‡ These authors contributed equally to this work

The antimicrobial defence of *Drosophila* relies largely on the challenge-induced synthesis of an array of potent antimicrobial peptides by the fat body^{1,2}. The defence against Gram-positive bacteria and natural fungal infections is mediated by the Toll signalling pathway, whereas defence against Gram-negative bac-

teria is dependent on the Immune deficiency (IMD) pathway^{3–18}. Loss-of-function mutations in either pathway reduce the resistance to corresponding infections^{3,9}. The link between microbial infections and activation of these two pathways has remained elusive. The Toll pathway is activated by Gram-positive bacteria through a circulating Peptidoglycan recognition protein (PGRP-SA)⁶. PGRPs appear to be highly conserved from insects to mammals, and the *Drosophila* genome contains 13 members^{19–23}. Here we report a mutation in a gene coding for a putative transmembrane protein, *PGRP-LC*, which reduces survival to Gram-negative sepsis but has no effect on the response to Gram-positive bacteria or natural fungal infections. By genetic epistasis, we demonstrate that *PGRP-LC* acts upstream of the *imd* gene. The data on *PGRP-SA* with respect to the response to Gram-positive infections, together with the present report, indicate that the PGRP family has a principal role in sensing microbial infections in *Drosophila*.

Our initial results on the role of *PGRP-SA* in the host defence of *Drosophila* prompted us to screen for mutations in other members of this family of genes. From a modified P transposon (XP) insertion library, we isolated a mutant line, *PGRP-LC*⁷⁴⁵⁴, containing a P-element inserted at nucleotide 72 in the first non-translated exon of the *PGRP-LC* gene (Fig. 1). This gene codes for a predicted 611 amino-acid protein of the *Drosophila* PGRP family. The predicted protein has no signal peptide²¹ and contains a non-canonical transmembrane domain (residues 275–294). The putative extracellular, carboxy-terminal region has three PGRP domains²³, one of which shows homology with *N*-acetylmuramoyl-L-alanine amidase. The predicted intracellular region (residues 1–275) has no significant homology to established protein domains.

When challenged with the Gram-negative bacteria *Escherichia coli* or *Agrobacterium tumefaciens*, the mutant flies showed a severely compromised induction in the levels of expression of the antibacterial peptide genes *Diptericin* (*Dpt*), *Attacin* and *Cecropin* (Fig. 2a). Induction of expression of the antifungal peptide gene *Drosomycin* (*Drs*) by Gram-positive or natural fungal infections was at wild-type levels in these mutant flies (Fig. 2b, c). This result contrasts markedly with our previous data on the *PGRP-SA* mutant *semmelweis* (*seml*) flies, in which induction of *Drs* expression caused by Gram-positive bacteria is blocked, whereas that of *Dpt* caused by Gram-negative bacteria is maintained at wild-type levels⁶ (Fig. 2).

Survival experiments with *PGRP-LC*⁷⁴⁵⁴ mutant flies showed a marked susceptibility to Gram-negative infections (*Agrobacterium*

tumefaciens, *Erwinia carotovora carotovora*, *Enterobacter cloacae*) and close to wild-type resistance to Gram-positive bacterial and natural fungal infections (Fig. 3). Again, this result contrasts with that of the *PGRP-SA* mutants, which are susceptible to Gram-positive infection but resistant to Gram-negative bacteria.

The results on the induction of antimicrobial peptides and the survival studies both indicate that *PGRP-LC* mutants have an immune response similar to that described for loss-of-function mutants of the IMD pathway. Overexpressing the *imd* gene has recently been reported to lead to challenge-independent induction of *Dpt* and other antibacterial peptide genes, but not to that of the antifungal peptide gene *Drs*¹⁸. We therefore examined the effects of overexpressing the *PGRP-LC* complementary DNA by generating larvae that contained a heat-shock promoter and Gal4 fusion gene (*hs-GAL4*) and a *UAS-PGRP-LC* transgene in a *Dpt-LacZ* and *Drs*-green fluorescent protein (*GFP*) reporter genetic background. After heat-shock treatment, and in the absence of any immune challenge, we observed strong expression in cells of the fat body of the *Dpt* but not the *Drs* reporter (Fig. 4a and data not shown). *PGRP-LC* can also be overexpressed in the *PGRP-LC*⁷⁴⁵⁴ insertion line under the control of a *UAS* site present in the transposon (Fig. 1), and we observed that in heterozygous *PGRP-LC*⁷⁴⁵⁴/*hs-GAL4* flies, *Dpt* was constitutively expressed (data not shown). We further used the fat-body-specific *yolk-GAL4* driver in place of the ubiquitous *hs-GAL4* driver. In these experiments, a strong constitutive expression of *Dpt* was detectable by quantitative reverse transcription polymerase chain reaction (RT-PCR) in extracts of adult females (Fig. 4b). Parallel experiments with male flies yielded no constitutive expression of the *Dpt* gene, consistent with the absence of Yolk protein in males. This result indicates that overexpression of *PGRP-LC* in the fat body is sufficient to activate expression of the *Dpt* gene.

The data presented so far clearly point towards an involvement of

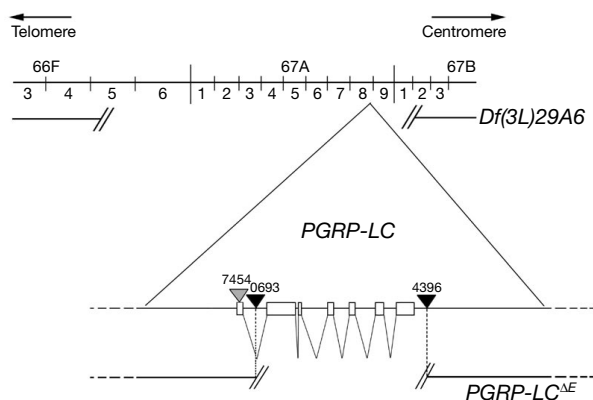


Figure 1 Schematic representation of the *PGRP-LC* locus. The *PGRP-LC* gene is located in 67A8, a genomic region uncovered by the deficiency *Df(3L)29A6*. 7454 is an XP transposable element inserted in the first exon of the *PGRP-LC* gene. Using two other XP transposable elements, 0693 and 4396, we generated a small deletion called *PGRP-LC*^{ΔE}, which removed the entire *PGRP-LC* coding region. The XP element contains two *UAS* sites in opposing directions at both ends of the transposon.

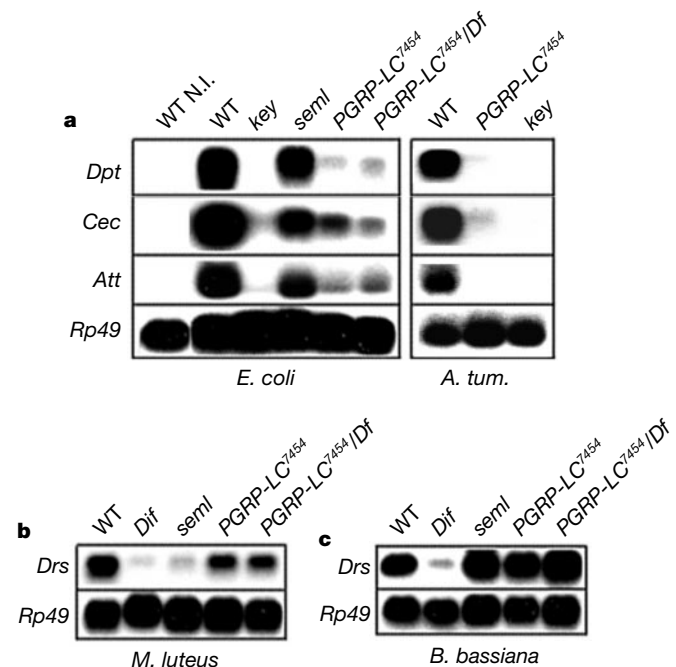


Figure 2 Expression of antimicrobial peptides in different mutant backgrounds after infection by Gram-negative bacteria, Gram-positive bacteria or fungi. Northern blots were performed with total RNA from wild-type (WT) flies, *seml*, *PGRP-LC*⁷⁴⁵⁴ mutant flies, or flies mutant for genes in the Toll signalling pathway (*Dif*) and in the IMD pathway (*key*). **a–c**, Flies were infected with *E. coli* and *A. tumefaciens* (**a**), *M. luteus* (**b**), or *B. bassiana* (**c**), as described in Methods. *Rp49* is used as an RNA-loading control. N.I., non-induced; *Drs*, *Drosomycin*; *Dpt*, *Diptericin*; *Att*, *Attacin*; *Cec*, *Cecropin*; *Df*, *Df(3L)29A6*.

the *PGRP-LC* gene in the activation of the IMD pathway. This pathway has been reported to induce cleavage of the latent cytoplasmic NF- κ B member Relish²⁴, allowing for nuclear translocation of the amino-terminal Rel homology domain fragment and activation of antibacterial peptide genes. Seven genes have been identified to date in the IMD–Relish intracellular signalling cascade, of which the most upstream gene is that coding for the death domain protein IMD^{9–18}. We performed genetic epistasis between the *PGRP-LC* and the *imd* genes using a *imd*^{shadok} null allele (D.F. *et al.*, manuscript in preparation), and observed that the constitutive activation of *Dpt* in the *PGRP-LC*⁷⁴⁵⁴/*hs-GAL4* flies was abolished in the *imd* mutant background (Fig. 4c). This result places the *PGRP-LC* gene upstream of *imd*. The mechanism by which overexpression of *PGRP-LC* activates the IMD pathway remains to be established. One possibility is that such an overexpression leads to the forced multimerization of a receptor complex that would bring in to close contact adapter proteins, hence initiating signalling.

Finally, to demonstrate that the *PGRP-LC*⁷⁴⁵⁴ mutant phenotype is indeed the result of the transposon insertion, we mobilized this transposon with a P-element transposase. We recovered several revertant lines that showed wild-type function with respect to *Dpt* inducibility and survival to *E. cloacae* infection (data not shown). As the overexpression of *PGRP-LC* from the *PGRP-LC*⁷⁴⁵⁴ transposon

can be obtained with a *hs-GAL4* driver (see above), we tested whether this overexpression is sufficient to rescue the *PGRP-LC*⁷⁴⁵⁴/*PGRP-LC*⁷⁴⁵⁴ homozygous mutant phenotype. To this end, we generated *hs-GAL4; PGRP-LC*⁷⁴⁵⁴/*PGRP-LC*⁷⁴⁵⁴ flies, which after heat shock became as resistant to *E. cloacae* infection as wild-type flies, in contrast to *PGRP-LC*⁷⁴⁵⁴/*PGRP-LC*⁷⁴⁵⁴ mutant flies (Fig. 4d).

The data presented above show conclusively that the full activation of the IMD pathway requires the presence of a wild-type *PGRP-LC* gene. However, in our experiments we consistently observed that the phenotypes (induction of antibacterial peptide genes, survival to Gram-negative infection) were less drastic than those reported for loss-of-function mutants of the *kenny* gene, which codes for the *Drosophila* homologue of IKK β /NEMO¹² (compare Figs 2a and 3) active in the IMD–Relish pathway. To rule out that this discrepancy reflected a possible hypomorphic character of the *PGRP-LC*⁷⁴⁵⁴ mutant line, we engineered a null mutation (*PGRP-LC* ^{ΔE}) by promoting the recombination between two additional XP elements flanking the *PGRP-LC* locus, thus generating a deletion of 8.7 kilobases (kb) of genomic DNA that encompasses the full coding region of the *PGRP-LC* gene (Fig. 1). Induction of *Dpt* and survival curves to infections in these null mutants were similar to those of the *PGRP-LC*⁷⁴⁵⁴ lines (data not shown and Fig. 3, *A. tumefaciens* graph). We conclude from these sets of data that *PGRP-LC* is not the sole upstream element

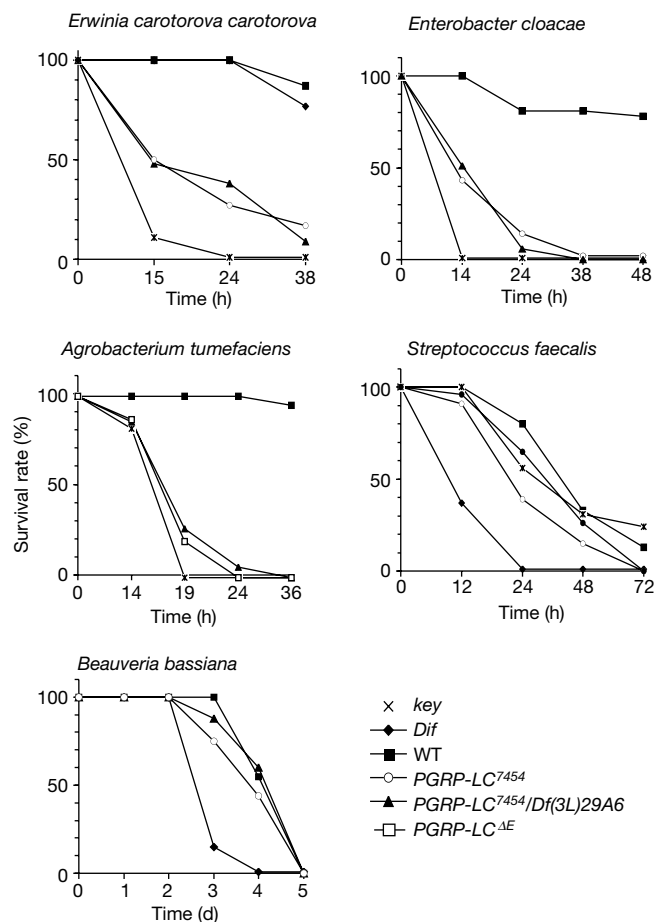


Figure 3 *PGRP-LC* mutant flies are highly susceptible to infection by Gram-negative bacteria. The survival rates of wild-type (WT), *Dif*², *key*¹, *PGRP-LC*⁷⁴⁵⁴, *PGRP-LC*⁷⁴⁵⁴/*Df(3L)29A6* and *PGRP-LC* ^{ΔE} flies infected with the indicated microorganisms are presented. Each experiment was performed using 25 flies per genotype and the results shown are representative of three independent experiments. *PGRP-LC*⁷⁴⁵⁴/*PGRP-LC*⁷⁴⁵⁴ homozygous mutant flies die as fast as *PGRP-LC*⁷⁴⁵⁴/*Df(3L)29A6* hemizygous mutant flies, indicating that this mutation is genetically null.

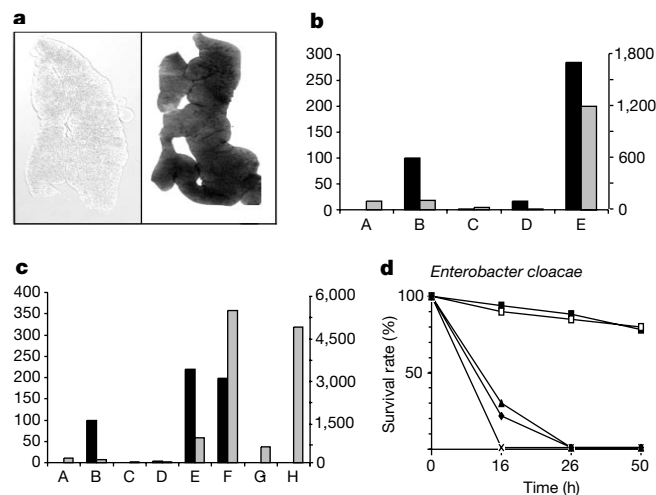


Figure 4 *PGRP-LC* is genetically upstream of *imd*. **a**, *Drs-GFP, Dpt-LacZ; hs-GAL4/+*; *UAS-PGRP-LC/+* third instar larvae were heat-shocked for 1 h at 37 °C. Twelve hours later, the fat body was dissected and stained for β -galactosidase activity and checked for GFP expression (not shown). *Drs-GFP, Dpt-LacZ; +/+*; *PGRP-LC*⁷⁴⁵⁴/*+* sibling larvae were used as a negative control (left panel). **b**, **c**, *Diapericin* (black bars) and *PGRP-LC* (grey bars) transcript levels were measured by quantitative RT-PCR. Results were first standardized against *Rp49* and then the value obtained for induced males or females was taken to be 100 as a reference. Left y-axis scale, *Diapericin*; right y-axis scale, *PGRP-LC*. **b**, Overexpression of *PGRP-LC*. A, non-challenged wild-type females; B, wild-type females challenged with *E. coli*; C, non-challenged *PGRP-LC*⁷⁴⁵⁴/*PGRP-LC*⁷⁴⁵⁴ females; D, challenged *PGRP-LC*⁷⁴⁵⁴/*PGRP-LC*⁷⁴⁵⁴ females; E, *yolk-GAL4/PGRP-LC*⁷⁴⁵⁴ non-challenged females. **c**, *imd* is epistatic to *PGRP-LC*. A, non-challenged wild-type males; B, challenged wild-type males (*E. coli*); C, non-challenged *PGRP-LC*⁷⁴⁵⁴/*PGRP-LC*⁷⁴⁵⁴ males; D, challenged *PGRP-LC*⁷⁴⁵⁴/*PGRP-LC*⁷⁴⁵⁴ males; E, non-heat-shocked *imd/+*; F, heat-shocked *imd/+*; G, *hs-GAL4/PGRP-LC*⁷⁴⁵⁴ non-challenged females; H, non-heat-shocked *imd/ind*; I, *hs-GAL4/PGRP-LC*⁷⁴⁵⁴ non-challenged females. **d**, Rescue of susceptibility to *E. cloacae* infection by *PGRP-LC* overexpression. Open squares, *hs-GAL4/+*; filled squares, *PGRP-LC*⁷⁴⁵⁴/*PGRP-LC*⁷⁴⁵⁴; filled triangles, non-heat-shocked *PGRP-LC*⁷⁴⁵⁴/*PGRP-LC*⁷⁴⁵⁴; filled diamonds, heat-shocked *PGRP-LC*⁷⁴⁵⁴/*PGRP-LC*⁷⁴⁵⁴; crosses, *key*.

activating the IMD pathway.

The present report together with our previous study⁶ indicates that Gram-positive and Gram-negative bacterial infections are sensed by distinct members of a diversified family of pattern recognition proteins. The name Peptidoglycan recognition protein was given to this family in reference to the initial reports of humoral proteins binding peptidoglycan in the moth *Trichoplusia ni*²⁰, and in the silkworm *Bombyx mori* where they activate a defence cascade leading to melanin formation^{19,25}. Adequate binding studies of PGRP family members to various microorganisms have not yet been performed; however, our data that PGRP-LC mediates the response to Gram-negative infection (this work) and that PGRP-SA is involved in the host defence against Gram-positive bacteria⁶, suggest that recognition by these proteins is not restricted to peptidoglycan patterns. We propose that the various *Drosophila* PGRP genes (and corresponding splice isoforms) can provide a recognition repertoire for microbial patterns that will detect many, if not most, of the microorganisms that can invade *Drosophila*. Mammals, like insects, synthesize both soluble (PGRP-S) and membrane-bound (PGRP-L, -I α , -I β) PGRPs²³. The human transmembrane PGRP-L protein is expressed in the liver, a functional homologue of the insect fat body, and binds to peptidoglycan and Gram-positive bacteria. Its functional relevance *in vivo* has not yet been reported. An unanswered question at present is how PGRP-LC signals to the IMD pathway. The putative intracytoplasmic region of PGRP-LC has no apparent sequence characteristics of intracellular signalling potential. We leave open the possibility that PGRP-LC acts as a co-receptor in the membrane of the fat body cell, and that activation of the death domain protein IMD during Gram-negative infection requires additional unidentified protein(s).

The field of study concerning the recognition of microbial pathogens in innate immunity has experienced a marked progression, both in insects and mammals, in the last few years. Whereas most studies to date have focused on Toll/Toll-like receptors^{26–28}, PGRPs are now coming to the forefront. In *Drosophila*, their functions as putative pattern recognition receptors appear to overshadow those of Tolls: the only Toll family member whose role in the immune response is clearly documented—Toll itself—requires activation through a PGRP sensing microbial infection⁶. □

Methods

Microbial strains

We used the following microbial organisms in this study: *Agrobacterium tumefaciens*, *Erwinia carotovora carotovora*, *Enterobacter cloacae*, *Escherichia coli* 1106 (Gram-negative bacteria); *Micrococcus luteus*, *Streptococcus faecalis* (Gram-positive bacteria); *Beauveria bassiana* (fungus).

Fly strains

Stocks were raised on standard cornmeal-agar medium at 25 °C. We used *white* and *yw* *Drs-GFP Dpt-LacZ*²⁹ flies as controls. *Dif*², *key*¹, *yw Drs-GFP Dpt-LacZ* flies were described elsewhere^{5,12,29}. The *Df(3L)29A* stock was obtained from the Bloomington stock centre.

Septic injury and survival experiments

Bacterial infections were performed by challenging adult flies with a thin tungsten needle previously dipped into a concentrated culture of the appropriate bacterial strains. To test antifungal response, anaesthetized flies were manually shaken for 30 s on a Petri dish containing a sporulating *Beauveria bassiana* culture³⁰. Infected animals were incubated at 25 °C (bacteria) or at 29 °C (fungus). Survival experiments were carried out with 25 flies for each tested genotype. Surviving flies were transferred daily into fresh vials and counted. Results are expressed as percentage of infected flies at different time points after infection. Each experiment is representative of at least three independent experiments. For northern blot analysis, flies were collected 24 h after infection and stored at –80 °C. For rescue experiments, flies were heat-shocked for 30 min at 37 °C. After 6 h incubation at 25 °C, flies were infected as described above.

Cloning and transformation of UAS-PGRP-LC

The wild-type PGRP-LC cDNA was obtained by PCR using expressed-sequence-tag (EST) clone (SD26236, ResGen) as a template. *EcoRI* and *XbaI* sites were introduced 5' and 3' respectively of the PGRP cDNA using the primers 5'-GGAATTCACAAATGCCTTTTGAACAATGAAACG-3' and 5'-GCTCTAGAGCTACGACCAATGAGTCCAGTT-3', and

the Platinum high fidelity Taq polymerase (Life Technologies). This fragment was then subcloned into pUAST. After sequencing, the construct was injected into *w¹¹¹⁸* embryos.

Quantitative RT-PCR

Samples of five flies were frozen in liquid nitrogen and crushed with a Mixer Mill 300 (Retch) twice for 50 s at 19 Hz. Total RNA was then prepared using the RNeasy kit (Qiagen) and RNAs eluted in 100 μ l RNase free water. Five microlitres were then used in a reverse transcription reaction using Superscript II reverse transcriptase (Invitrogen) and random primers (Invitrogen) according to the supplier's instructions. The cDNAs were then diluted to a proper concentration so that the subsequent PCR reactions would not be inhibited by components of the reverse transcription preparation. PCR reactions were set up using the qPCR kit (Eurogentec) and in a 1/75,000 final concentration of SYBRGreen. Real-time PCR was then performed in 96-well plates on an i-cycler iQ (Biorad)—usual PCR conditions of pre-incubation at 95 °C, 40 cycles; 15 s at 95 °C, 1 min at 60 °C. PCR reactions were done in duplicates; to check for specificity of the PCR reaction, melting curves were analysed for each data point. The level of expression of the gene of interest was then normalized against the measured level of the RNA coding for Ribosomal protein 49 determined in each sample. Primers were as follows: *Diptericin*, forward 5'-GCTGCGC AATCGCTTCTACT-3', reverse 5'-TGGTGGAGTGGGCTTCATG-3'; *Rp49*, forward 5'-GACGCTTCAAGGACAGTATCTG-3', reverse 5'-AAACGCGGTTCTGCATGAG-3'; PGRP-LC, forward 5'-GGCAGTTCGAATCGAAATCG-3', reverse 5'-CGCTGATACGCTT GGATTCC-3'.

Received 22 February; accepted 5 March 2002.

Published online 24 March 2002, DOI 10.1038/nature734.

- Hoffmann, J. A. & Reichhart, J. M. *Drosophila* innate immunity: an evolutionary perspective. *Nature Immunol.* **3**, 121–126 (2002).
- Tzou, P., De Gregorio, E. & Lemaitre, B. How *Drosophila* combats microbial infection: a model to study innate immunity and host–pathogen interactions. *Curr. Opin. Microbiol.* **5**, 102–110 (2002).
- Lemaitre, B., Nicolas, E., Michaut, L., Reichhart, J. M. & Hoffmann, J. A. The dorsoventral regulatory gene cassette *spätzle/Toll/cactus* controls the potent antifungal response in *Drosophila* adults. *Cell* **86**, 973–983 (1996).
- Meng, X., Khanuja, B. S. & Ip, Y. T. Toll receptor-mediated *Drosophila* immune response requires Dif, an NF- κ B factor. *Genes Dev.* **13**, 792–797 (1999).
- Rutschmann, S. *et al.* The Rel protein DIF mediates the antifungal, but not the antibacterial, response in *Drosophila*. *Immunity* **12**, 569–580 (2000).
- Michel, T., Reichhart, J., Hoffmann, J. A. & Royet, J. *Drosophila* Toll is activated by Gram-positive bacteria through a circulating peptidoglycan recognition protein. *Nature* **414**, 756–759 (2001).
- Tauszig-Delamasure, S., Bilak, H., Capovilla, M., Hoffmann, J. A. & Imler, J. L. *Drosophila* MyD88 is required for the response to fungal and Gram-positive bacterial infections. *Nature Immunol.* **3**, 91–97 (2002).
- Rutschmann, S., Kilinc, A. & Ferrandon, D. The Toll pathway is required for resistance to Gram-positive bacterial infections in *Drosophila*. *J. Immunol.* **168**, 1542–1546 (2002).
- Lemaitre, B. *et al.* A recessive mutation, *immune deficiency (imd)*, defines two distinct control pathways in the *Drosophila* host defence. *Proc. Natl Acad. Sci. USA* **92**, 9465–9469 (1995).
- Hedengren, M. *et al.* *Relish*, a central factor in the control of humoral but not cellular immunity in *Drosophila*. *Mol. Cell* **4**, 1–20 (1999).
- Elrod-Erickson, M., Mishra, S. & Schneider, D. Interactions between the cellular and humoral immune responses in *Drosophila*. *Curr. Biol.* **10**, 781–784 (2000).
- Rutschmann, S. *et al.* Role of *Drosophila* IKK γ in a Toll-independent antibacterial immune response. *Nature Immunol.* **1**, 342–347 (2000).
- Silverman, N. *et al.* A *Drosophila* I κ B kinase complex required for Relish cleavage and antibacterial immunity. *Genes Dev.* **14**, 2461–2471 (2000).
- Leulier, F., Rodriguez, A., Khush, R. S., Abrams, J. M. & Lemaitre, B. The *Drosophila* caspase Dredd is required to resist Gram-negative bacterial infections. *EMBO Rep.* **1**, 353–358 (2000).
- Lu, Y., Wu, L. P. & Anderson, K. V. The antibacterial arm of the *Drosophila* innate immune response requires an I κ B kinase. *Genes Dev.* **15**, 104–110 (2001).
- Vidal, S. *et al.* Mutations in the *Drosophila* *dTAK1* gene reveal a conserved function for MAPKKs in the control of rel/NF- κ B dependent innate immune responses. *Genes Dev.* **15**, 1900–1912 (2001).
- Wu, L. P., Choe, K. M., Lu, Y. & Anderson, K. V. *Drosophila* immunity: genes on the third chromosome required for the response to bacterial infection. *Genetics* **159**, 189–199 (2001).
- Georgel, P. *et al.* *Drosophila* Immune Deficiency (IMD) is a Death Domain protein that activates antibacterial defense and can promote apoptosis. *Dev. Cell* **1**, 503–514 (2001).
- Yoshida, H., Kinoshita, K. & Ashida, M. Purification of a peptidoglycan recognition protein from hemolymph of the silkworm, *Bombyx mori*. *J. Biol. Chem.* **271**, 13854–13860 (1996).
- Kang, D., Liu, G., Lundstrom, A., Gelius, E. & Steiner, H. A peptidoglycan recognition protein in innate immunity conserved from insects to humans. *Proc. Natl Acad. Sci. USA* **95**, 10078–10082 (1998).
- Werner, T. *et al.* A family of peptidoglycan recognition proteins in the fruit fly *Drosophila melanogaster*. *Proc. Natl Acad. Sci. USA* **97**, 13772–13777 (2000).
- Liu, C., Gelius, E., Liu, G., Steiner, H. & Dziarski, R. Mammalian peptidoglycan recognition protein binds peptidoglycan with high affinity, is expressed in neutrophils, and inhibits bacterial growth. *J. Biol. Chem.* **275**, 24490–24499 (2000).
- Liu, C., Xu, Z., Gupta, D. & Dziarski, R. Peptidoglycan recognition proteins: a novel family of four human innate immunity pattern recognition molecules. *J. Biol. Chem.* **276**, 34686–34694 (2001).
- Stöven, S., Ando, I., Kadalayil, L., Engström, Y. & Hultmark, D. Activation of the *Drosophila* NF- κ B factor Relish by rapid endoproteolytic cleavage. *EMBO Rep.* **1**, 347–352 (2000).
- Ochiai, M. & Ashida, M. A pattern recognition protein for peptidoglycan. Cloning the cDNA and the gene of the silkworm, *Bombyx mori*. *J. Biol. Chem.* **274**, 11854–11858 (1999).
- Sieling, P. A. & Modlin, R. L. Toll-like receptors: mammalian 'taste receptors' for a smorgasbord of microbial invaders. *Curr. Opin. Microbiol.* **5**, 70–75 (2002).
- Akira, S., Takeda, K. & Kaisho, T. Toll-like receptors: critical proteins linking innate and acquired

- immunity. *Nature Immunol.* **2**, 675–680 (2001).
28. Kimbrell, D. A. & Beutler, B. The evolution and genetics of innate immunity. *Nature Rev. Genet.* **2**, 256–267 (2001).
29. Jung, A., Crikui, M.-C., Rutschmann, S., Hoffmann, J.-A. & Ferrandon, D. A microfluorometer assay to measure the expression of β -galactosidase and GFP reporter genes in single *Drosophila* flies. *Biotechniques* **30**, 594–601 (2001).
30. Lemaitre, B., Reichhart, J. M. & Hoffmann, J. A. *Drosophila* host defense: differential display of antimicrobial peptide genes after infection by various classes of microorganisms. *Proc. Natl Acad. Sci. USA* **94**, 14614–14619 (1997).

Acknowledgements

We thank our colleagues in the laboratory for critical comments on the manuscript; L. Troxler and C. Hetru for computer analysis; and M. E. Moritz and M. Schneider for providing bacterial and fungal cultures. This work was supported by CNRS, the Ministère de l'Éducation Nationale de la Recherche et de la Technologie and the Fondation pour la Recherche Médicale (Implantation jeunes équipes to J.R. and D.F.). Financial support from the National Institutes of Health is acknowledged.

Competing interests statement

The authors declare competing financial interests: details accompany the paper on *Nature's* website (<http://www.nature.com>).

Correspondence and requests for materials should be addressed to D.F. (e-mail: D.Ferrandon@ibmc.u-strasbg.fr) or J. R. (e-mail: J.Royet@ibmc.u-strasbg.fr).

Functional genomic analysis of phagocytosis and identification of a *Drosophila* receptor for *E. coli*

Mika Rämet^{*,†}, Pascal Manfrulli^{*,†}, Alan Pearson^{*}, Bernard Mathey-Prevot[†] & R. Alan B. Ezekowitz^{*}

^{*} Laboratory of Developmental Immunology, Massachusetts General Hospital for Children, and Department of Pediatrics, Harvard Medical School, 55 Fruit Street, Boston, Massachusetts 02114, USA

[†] Division of Pediatric Oncology, Dana-Farber Cancer Institute, Harvard Department of Pediatrics, 44 Binney Street, Boston, Massachusetts 02115, USA

[†] These authors contributed equally to this work

The recognition and phagocytosis of microbes by macrophages is a principal aspect of innate immunity that is conserved from insects to humans. *Drosophila melanogaster* has circulating macrophages that phagocytose microbes similarly to mammalian macrophages^{1,2}, suggesting that insect macrophages can be used as a model to study cell-mediated innate immunity. We devised a double-stranded RNA interference-based screen in macrophage-like *Drosophila* S2 cells, and have defined 34 gene products involved in phagocytosis. These include proteins that participate in haemocyte development, vesicle transport, actin cytoskeleton regulation and a cell surface receptor. This receptor, Peptidoglycan recognition protein LC (PGRP-LC), is involved in phagocytosis of Gram-negative but not Gram-positive bacteria. *Drosophila* humoral immunity also distinguishes between Gram-negative and Gram-positive bacteria through the Imd and Toll pathways, respectively; however, a receptor for the Imd pathway has not been identified. Here we show that PGRP-LC is important for antibacterial peptide synthesis induced by *Escherichia coli* both *in vitro* and *in vivo*. Furthermore, *totem* mutants, which fail to express PGRP-LC, are susceptible to Gram-negative (*E. coli*), but not Gram-positive, bacterial infection. Our results demonstrate that PGRP-LC is an essential component for recognition and signalling of Gram-negative bacteria. Furthermore, this functional genomic approach is likely to have appli-

cations beyond phagocytosis.

The ligation of phagocytic receptors by microbes triggers a reorganization of the plasma membrane and cytoskeletal components that facilitates the ingestion of the bound particle and the formation of a phagosome³. A recent analysis identified greater than 160 phagosome-associated proteins in mouse macrophages⁴, demonstrating the complexity of the phagosome and phagocytosis. The functional significance of these and other proteins for phagocytosis remains undefined in many instances. This complexity led us to seek a potentially simpler model system to study phagocytosis. We and others have demonstrated that *Drosophila* macrophages express phagocytic properties that are similar to mammalian macrophages^{1,2}. We have also shown that phagocytosis is similar in *Drosophila* S2 cells and primary *Drosophila* macrophages⁵, so that S2 cells are a valid *in vitro* system for studying phagocytosis. Furthermore, double-stranded RNA interference (RNAi) specifically and effectively silences the expression of targeted genes in S2 cells⁶, thereby providing an opportunity to define genes whose products have a function in phagocytosis *in vitro*, allowing us to use this information to specify *in vivo* experiments.

Double-stranded RNAs (dsRNAs) were synthesized using ran-

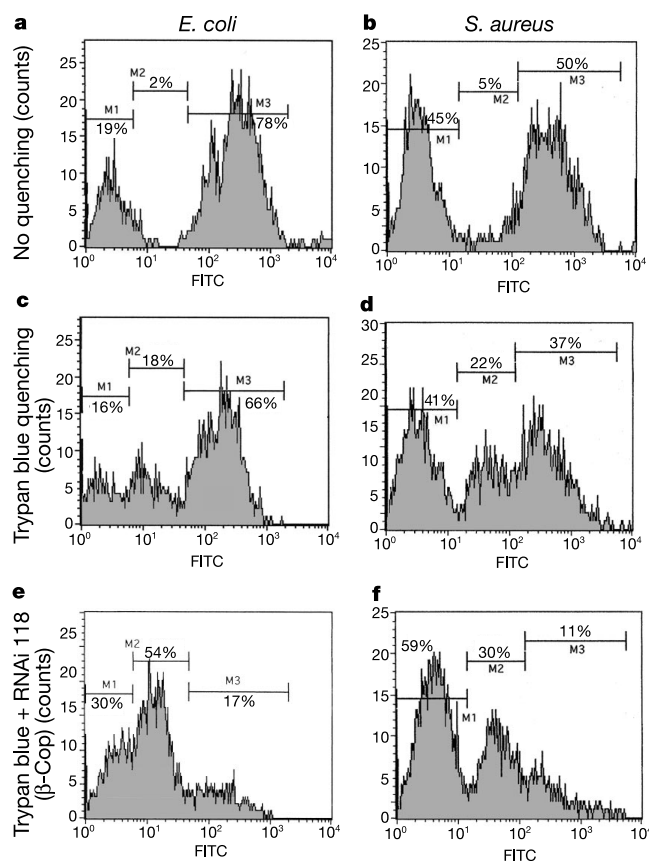


Figure 1 Quantification of phagocytosis of *E. coli* (a, c, e) and *S. aureus* (b, d, f) by FACS. S2 cells were incubated with FITC-labelled bacteria for 30 min at 4 °C followed by 15 (*E. coli*) or 20 (*S. aureus*) min incubation at 26 °C to allow internalization. Gate M1 represents the background autofluorescence of S2 cells; gate M3 in a, b, represents brightly fluorescent S2 cells that have bound and/or internalized FITC-labelled bacteria. c, d, Representative FACS scans after Trypan blue quenching. The M3 peak represents S2 cells that have internalized FITC-labelled bacteria, and the M2 peak represents S2 cells that have bound but not internalized FITC-labelled bacteria. e, f, RNAi treatment (β -Cop) causing decreased phagocytosis. The amount of phagocytosis was quantified as the percentage of cells within the gate M3 multiplied by the mean fluorescence intensity of these cells.

dom templates from an S2-cell-derived complementary DNA library⁷. Two independent dsRNAs were pooled and incubated with 5.0×10^5 S2 cells for 60 h. Viability and gross morphology of these cells were assayed by phase contrast microscopy and Trypan blue exclusion. The phagocytic ability of treated cells was compared with untreated cells. Figure 1a, b shows typical fluorescence-activated cell sorting (FACS) scans of S2 cells incubated with heat-killed, fluorescein isothiocyanate (FITC)-labelled *E. coli* or *Staphylococcus aureus* for 30 min at 4 °C, followed by 15 or 20 min incubation at 26 °C to allow internalization. Gate M3 in Fig. 1a, b shows the S2 cell population that has bound and/or internalized FITC-labelled bacteria. Gate M1 represents the background autofluorescence of S2 cells that have no associated FITC-labelled microbes. To distinguish between bound and internalized bacteria we incubated the cells with Trypan blue just before FACS analysis to quench most of the fluorescence of bound particles, but not of particles that have been internalized^{8,9}. Figure 1c, d shows representative FACS scans after Trypan blue quenching. There is a shift in fluorescence from the bright M3 peak to the dim M2 peak (compare Fig. 1a with c and b with d). The M3 peak represents the S2 cells that have bright fluorescence (no quenching of Trypan blue) representing internalized bacteria, whereas the M2 peak represents S2 cells that have bound but not internalized FITC-labelled bacteria.

We examined 1,000 random dsRNAs for their effect on S2 cell

phagocytosis of FITC-labelled *E. coli* and *S. aureus*. Most had little or no effect. 56 RNAi treatments representing 45 different genes showed a detectable phenotype as defined by increased number of dead cells, morphological changes, or at least a 20% reduction in phagocytosis. As shown in Supplementary Information Table 1, 15 RNAi treatments representing nine genes caused increased cell death. These included apparently essential genes coding for transcription factors, proteins required for RNA processing, DNA replication, ubiquitination and lipid metabolism. A total of 38 RNAi treatments representing 34 genes resulted in at least a 20% decrease in phagocytosis without a significant increase in cell death (Supplementary Information Table 1). An example of an RNAi treatment that yielded a phenotype of strongly reduced phagocytosis is shown in Fig. 1e, f. Sequence analysis of the template plasmid (designated 118) indicated that the targeted gene codes for vesicle coat protein β -Cop, a constituent of the COPI complex². A role for COPI in mammalian crystallizable fragment (Fc) receptor-mediated phagocytosis has recently been postulated¹⁰. In one study, defective phagocytosis of COPI-deficient cells was attributed in part to decreased vesicular membrane sources required for pseudopod extension¹⁰. Of note, two other RNAi treatments (designated 221 and 929) targeting β -Cop gene expression independently gave the same phenotype, indicating that the results from the screen are reproducible. Furthermore, RNAi treatment targeting a gene coding for another component of the COPI complex, δ -Cop⁹, was also identified (Supplementary Information Table 1). In addition, three RNAi treatments targeted gene products involved in clathrin-coated vesicle formation, resulting in moderately decreased phagocytosis (Supplementary Information Table 1). These results indicate that *Drosophila* phagocytosis resembles mammalian phagocytosis^{10,11} and that vesicular transport is important for efficient phagocytosis.

Our screen identified four RNAi treatments that targeted conserved haematopoietic transcription factors including Serpent, Nejire, Myb and Pcaf (Supplementary Information Table 1). In embryos, there are two principal classes of *Drosophila* haematopoietic cells: plasmocytes/macrophages and crystal cells¹². The *Drosophila* GATA-factor Serpent is required for the development of both lineages, as it regulates the expression of the crystal-cell-determining factor *lozenge* and plasmatocyte-determining factor *glial cells missing*^{12,13}. Inactivation of *serpent* in S2 cells resulted in a change in cell morphology, a slowing in cell growth without any apparent effect on cell viability, and an average 77% reduction in phagocytosis. It appears therefore that Serpent is not only a differentiation factor but may also have a role in regulating genes that are involved in constitutive macrophage functions such as phagocytosis. Polymerase chain reaction with reverse transcription (RT-PCR) analysis revealed that S2 cells express *glial cells missing* but not *lozenge* (data not shown), further supporting the concept that S2 cells are macrophage-like. Interestingly, RNAi treatment targeting the transcription factor Nejire, a *Drosophila* homologue of the mammalian Creb-binding protein (Cbp) that is necessary for haematopoiesis in the mouse¹⁴, decreased phagocytosis by approximately 50%.

As expected, we identified a number of structural and cytoskeletal regulatory proteins including Actin 5C (Supplementary Information Table 1). Two putative membrane to actin linking proteins were identified: Coracle, a *Drosophila* member of the protein 4.1 superfamily¹⁵, and Short stop, a protein related to Dystrophin that is expressed in axons during development¹⁶. Coracle seems to have an essential role in epithelial tight-junction formation and the maintenance of the trans-epithelial barrier. Our findings suggest that it has a role in macrophage phagocytosis as well. Short stop is involved in axon guidance, which—like the formation of filopodia and lamellipodia in phagocytosis—requires rearrangements of the actin cytoskeleton. The association between phagocytosis and axon guidance gains further credence by the identification of the

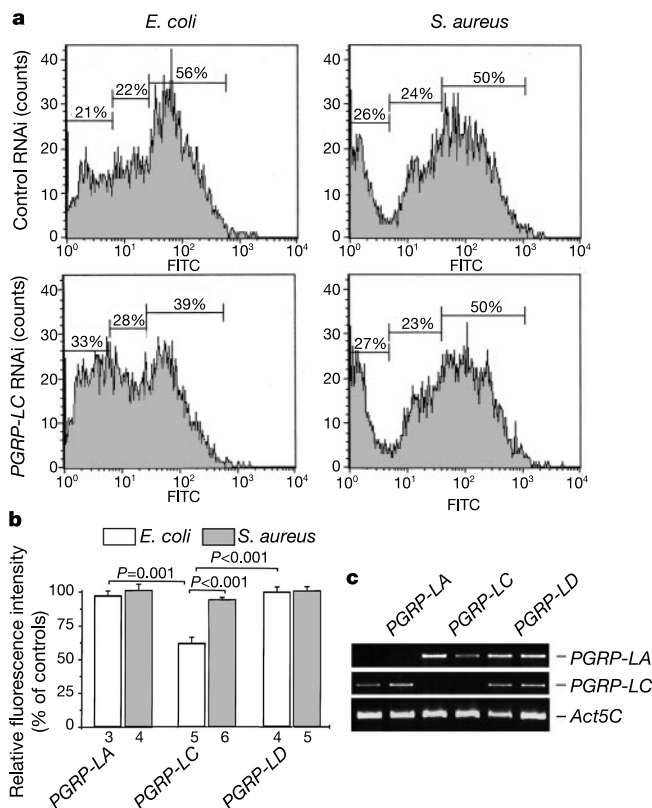


Figure 2 PGRP-LC RNAi decreases phagocytosis of *E. coli* but does not affect phagocytosis of *S. aureus*. **a**, FACS analysis of phagocytosis of heat-killed, FITC-labelled *E. coli* and *S. aureus* by S2 cells treated with PGRP-LC RNAi (lower panel) or with sham RNAi (upper panel). **b**, Effect of PGRP-LA, PGRP-LC and PGRP-LD RNAi on phagocytosis of FITC-labelled *E. coli* and *S. aureus* as quantified by FACS. Results show the level of phagocytosis compared with control RNAi (ds47)-treated S2 cells. The number (*n*) of independent experiments is shown below each column. Error bars represent standard error. Data were analysed using one-way ANOVA. $P < 0.05$ was considered to be significant. **c**, PGRP-LA (first two lanes) and PGRP-LC (middle lanes) RNAi treatments specifically decrease RNA levels of the targeted gene. A representative RT-PCR for PGRP-LA, PGRP-LC and Actin 5C is shown.

gene *longitudinals lacking (lola)* in our screen. Lola is a nuclear factor required for axon growth and guidance in the *Drosophila* embryo¹⁷.

We found two putative signalling molecules: Abelson interacting protein (Abi)¹⁸ and Tre oncogene-related protein (RN-tre). As *Abelson tyrosine kinase (Abl)*, *enabled* or *Disabled* dsRNA treatments

had no effect on phagocytosis in S2 cells (data not shown), the function of Abi in phagocytosis appears to be independent of Abl. The effectiveness of *Abl* RNAi was confirmed by semi-quantitative RT-PCR (data not shown). RN-tre is a Rab5 GTPase-activating protein¹⁹ that may affect Rab5-regulated vesicle movements.

In our screen we found one RNAi treatment that reduced

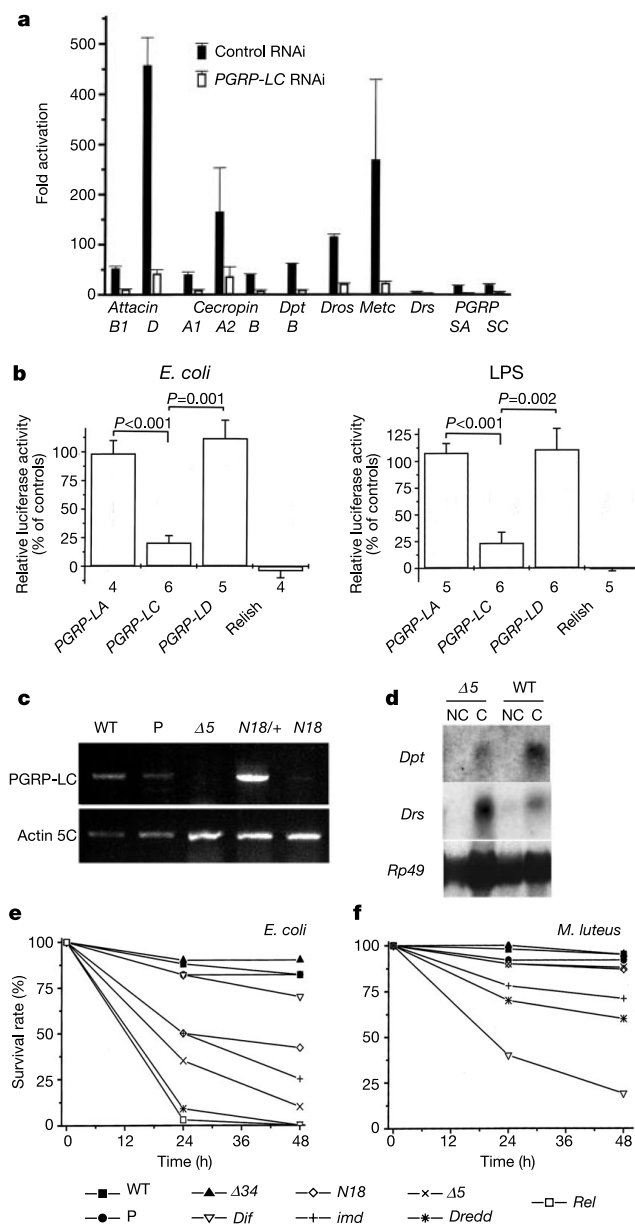


Figure 3 *PGRP-LC* controls Imd pathway-mediated response to Gram-negative bacteria both *in vitro* and *in vivo*. **a**, Synthesis of antimicrobial peptides induced by *E. coli* is blocked by *PGRP-LC* RNAi. Oligonucleotide microarrays were used to obtain genome-wide analysis of genes affected by 6-h exposure to *E. coli*. All genes upregulated by >20-fold in control RNAi (*ds47*)-treated cells (black columns) in response to *E. coli* compared with control RNAi-treated S2 cells without exposure to bacteria, are shown. *PGRP-LC* RNAi effectively inhibits this induction (white columns). Data are mean of two comparisons $\pm$ standard error. Two *Attacin* (*Att*) (B1 and D) and three *Cecropin* genes (A1, A2, B) were induced. *Dpt*, *Diptericin*; *Drs*, *Drosomycin*; *Dros*, *Drosocin*; *Metc*, *Metchnikowin*. **b**, *PGRP-LC* is required for *Att* expression in S2 cells. Cells were incubated for 12 h with *E. coli* or 10 μ M LPS, and thereafter *Att* reporter activity was measured using a luciferase assay. *Actin 5C* promoter-driven β -galactosidase expression was used to normalize the results. Results are shown as the percentage of luciferase induction compared with control RNAi (*ds47*)-treated cells. The number (*n*) of independent RNAi treatments from four

independent experiments is shown below each column. On average 5.7 ± 1.9 fold induction was observed with *E. coli* and 3.5 ± 0.4 fold induction with LPS. Error bars represent standard error. One-way ANOVA was used to determine statistical significance. **c**, RT-PCR analysis of *PGRP-LC* expression in different flies. WT, wild type; P, BG00650; $\Delta 5$, an imprecise excision of BG00650; *N18/+* and *N18*, heterozygous and homozygous flies for a new insertion of BG00650 in the first intron of *PGRP-LC*. **d**, Northern blot analysis of antimicrobial peptides 12 h after immune challenge with *E. coli* in $\Delta 5$ mutant and wild-type flies. NC, non-challenged; C, challenged. **e**, *totem* flies are susceptible to infection by *E. coli* bacteria. Survival rates (expressed in percentage) of wild-type (Oregon R), P (BG00650), $\Delta 34$ (a perfect excision of BG00650), *Dif* (*Dif*¹), *N18*, *imd* (*imd*¹), $\Delta 5$, *Dredd* (*Dredd*^{B118}) and *Rel* (*Relish*^{E28}) after infection with *E. coli* and *M. luteus*. Thirty flies of each line were pricked by a needle dipped into *E. coli* ($A_{600} = 150$) (**e**) or *M. luteus* ($A_{600} = 150$) (**f**).

phagocytosis of *E. coli* but did not affect phagocytosis of *S. aureus* (Fig. 2). This gene codes for the putative transmembrane protein PGRP-LC, which belongs to a family of *Drosophila* genes defined by their conserved peptidoglycan recognition domains²⁰. PGRPs were first identified from the moth *Trichoplusia ni*²¹ and the silkworm *Bombyx mori*²², but PGRP family members have also been identified in flies, mouse²³ and humans²⁴. Several of the mammalian, silkworm and moth PGRPs have been demonstrated to bind selectively Gram-positive bacteria. Recently, a loss-of-function mutation in the *Drosophila* PGRP-SA gene *semmelweis* that significantly decreased resistance to Gram-positive bacterial infection, but not to fungal or gram-negative bacterial infection, has been identified²⁵. Three of the *Drosophila* PGRPs are putative membrane proteins: PGRP-LC, -LA and -LD. Only PGRP-LC RNAi resulted in a reduction both in phagocytosis (Fig. 2b) and binding (data not shown) of *E. coli*. The specificity and effectiveness of the RNAi treatments were demonstrated by RT-PCR (Fig. 2c). These results suggest that PGRP-LC may function as a pattern recognition receptor that mediates the cell's response to *E. coli*. The partial decrease in phagocytosis after PGRP-LC RNAi probably reflects the participation of other receptors—such as dSR-CI—in the recognition of Gram-negative bacteria⁵.

To determine whether PGRP-LC has a function in signalling the antimicrobial peptide response in S2 cells, we performed a genome-wide analysis of genes induced by 6-h exposure to *E. coli* using oligonucleotide microarrays. Figure 3 shows the genes that were induced by greater than 20-fold in control RNAi-treated cells in response to *E. coli*, all of which are regulated by the Imd pathway and code for proteins with immunity-related functions. PGRP-LC RNAi markedly reduced the induction of all of these genes (Fig. 3a, white columns). For example the 400-fold induction of *Attacin* in control cells was markedly reduced by PGRP-LC RNAi. Toll-pathway-dependent *Drosomycin* was not induced by *E. coli* in S2 cells (Fig. 3a). To further validate the role for PGRP-LC in mediating cellular responses to *E. coli* and lipopolysaccharide (LPS), we used an *Attacin* promoter-driven luciferase reporter assay as described previously^{5,26}. RNAi targeting PGRP-LC, but not PGRP-LA or PGRP-LD, reduced luciferase induction in response to a 12-h incubation with *E. coli* or LPS by 80% in S2 cells (Fig. 3b). As expected, RNAi targeting the transcription factor *Relish* abolished *E. coli*-induced luciferase expression. Of note, *Drosophila* is relatively resistant to LPS, as high concentrations are required to evoke biological responses²⁷, and whether LPS *per se* or a contaminant in the preparation is the elicitor of the response when these high doses of LPS are used, remains unclear.

To test whether the function assigned to PGRP-LC *in vitro* was operative *in vivo*, we created mutants for PGRP-LC. We induced the excision of a modified P-element (GT1, line BG00650 (ref. 28)), which was inserted in the promoter region of PGRP-LC at 53 nucleotides from the first untranslated exon. We identified one deletion ($\Delta 5$) and a new insertion mutant in PGRP-LC (*N18*). $\Delta 5$ resulted from the imprecise excision of the P-element that abolished PGRP-LC expression as determined by RT-PCR (Fig. 3c, middle lane). PCR mapping of the region using specific complementary primers indicates that the entire first translated exon of PGRP-LC is deleted. Quantitative RT-PCR on *N18* flies revealed that PGRP-LC gene expression was greatly reduced (Fig. 3c, last lane). The expression of the two genes that flank the PGRP-LC locus (PGRP-LA and *CG4347*) appears similar to wild-type flies (data not shown). Taken together, we believe that we have identified two mutant alleles of PGRP-LC that we call *totem* (*tot*): $\Delta 5$ is a null allele and *N18* is an hypomorphic allele.

After challenge with the Gram-negative bacteria *E. coli*, we found that the accumulation of *Diptericin* messenger RNA was greatly reduced in $\Delta 5$ mutant flies (Fig. 3d). By contrast, these mutant flies showed evident *Drosomycin* mRNA expression. Similar results were obtained with *N18* mutant flies (results not shown). Survival

experiments with the two *tot* alleles were conducted after infection with *E. coli* and *Micrococcus luteus*. $\Delta 5$ flies are clearly susceptible to infection by *E. coli* bacteria compared with wild-type flies, flies that contain the original P-element insertion (BG00650), flies that have perfect excision of the P-element ($\Delta 34$), and *Dif* mutant flies (Fig. 3e). In addition, the $\Delta 5$ flies died earlier than the hypomorphic *imd* mutant flies, although later than *Dredd* and *Relish* mutant flies. Hypomorphic *tot* mutant flies, *N18*, also displayed enhanced susceptibility to *E. coli* infection. Importantly, *tot* flies were as resistant to infection with the Gram-positive bacteria *M. luteus* as wild-type flies, control flies (BG00650 and $\Delta 34$) and mutant flies in the *imd* pathway (*imd*¹ and *Dredd*^{B118}) (Fig. 3f).

We have developed a functional screen that has defined genes necessary for phagocytosis. The advantage of an *in vitro* system is that it can be used to rapidly screen for genes of interest, some which are embryonic lethal. The next step is to correlate the *in vitro* findings with mutants. Our results *in vitro* and *in vivo* suggest that PGRP-LC is a pattern recognition receptor in the Imd pathway. As such it appears that in *Drosophila* the PGRP-LC family of molecules are important bacterial recognition proteins. It also seems that the original description as peptidoglycan recognition molecules may be too restrictive, and that the range of pattern recognition should at least extend to *E. coli*. The presence of human PGRP homologues raises interesting questions as to the role of these molecules in mammalian innate immunity. Our results also demonstrate that *Drosophila* is a useful model system to study phagocytosis, as one can combine *in vitro* observations with the tool of *Drosophila* genetics. □

Methods

RNAi treatments

A cDNA library derived from S2 cells cloned into pcDNA1 plasmids⁷ was used as a source for templates for RNA synthesis. pcDNA1 has T7 and SP6 sequences that can be used to generate both sense and antisense RNA of the insert. First, *E. coli* bacteria transformed with the cDNA library were plated onto ampicillin (20 mg ml⁻¹) and tetracycline (8 mg ml⁻¹) containing plates. Next, plasmid DNAs isolated from individual colonies using Qiagen Mini columns (Qiagen) were used as templates for dsRNA synthesis. When specific RNAi treatments were performed, the anticipated gene was cloned into the pcDNA1 plasmid. The following primers were used: PGRP-LA, 5'-CCGTCGCGATTATCGGTCAA-3' and 5'-CTGAATGGTGGCGCATCTTGA-3'; PGRP-LC, 5'-CATGATGTGATGTTGCCGA-3' and 5'-TTGCTTGCCACAATGCCA-3'; and PGRP-LD, 5'-TCAGTTAGTGGAAGATGC-3' and 5'-TTGTCTCCAGAGCAAATAGG-3'. RNA synthesis was carried out from the linearized template using the T7 and SP6 MegaScript RNA polymerases (Ambion) essentially as recommended by the manufacturer. We used two plasmids per reaction as templates for RNA synthesis. After DNAase treatment, complementary RNAs were annealed and yield was determined using an absorbance of 260 nm (*A*₂₆₀). A total of 5.0 × 10⁵ S2 cells were treated with 40 µg dsRNA (20 µg per gene) for 60 h, and then analysed for viability and gross morphology. Subsequently, the ability of the RNAi-treated cells to phagocytose heat-killed, FITC-labelled *E. coli* and *S. aureus* bacteria (Molecular Probes) was compared with controls using flow cytometry.

FACS analysis to quantify phagocytosis

A total of 2.0 × 10⁵ RNAi-treated S2 cells were plated into 48-well plates and incubated for 2 h at 26 °C. Thereafter the cells were cooled down to 4 °C for 30 min, and 2.0 × 10⁶ heat-killed, FITC-labelled *E. coli* or *S. aureus* bacteria were added. Each RNAi-treated sample was done as duplicate for both *E. coli* and *S. aureus*. Plates were centrifuged for 1 min at 100g and incubated for 30 min at 4 °C to allow binding. Then cells were incubated at 26 °C for the indicated times. Thereafter plates were placed on ice and samples analysed with FACS using the CELLQuest program (Becton Dickinson). The fluorescence of extracellular particles was quenched by replacing the medium with 0.2% Trypan blue in 1 × PBS, pH 5.5, shortly before the actual measurement. S2 cells showing increased fluorescence were gated as shown in Fig. 1. The amount of phagocytosis was quantified as the percentage of fluorescence-positive cells within gate M3, multiplied by the mean fluorescence of these cells. We counted 3,000–10,000 cells from each sample.

RT-PCR analysis

Total RNA was isolated using either Trizol reagent (GIBCO BRL) or the Rneasy Mini Kit (Qiagen), and RT-PCR was carried out according to standard procedures with gene-specific primers. Primers were: PGRP-LA, 5'-GCCATTGCTGAGAAGGGTCT-3' and 5'-TGATAGAGGTGTGACTGCGGC-3'; PGRP-LC, 5'-TTGGCAGCATCGCCCTGACC-3' and 5'-CTTGGCTAGAGCGCATCCGC-3'; *Actin*, 5C 5'-CGAAGAAGTTGCTGCTCTGG-3' and 5'-AGAACGATACCGGTGGTACG-3'; and *Abl*, 5'-TGATAACAGCTGGTCAAGGT-3' and 5'-AGCTGCTTTCCACCGCTT-3'.

Analysis of mRNA expression using oligonucleotide microarrays

Total RNA was extracted from 1.0×10^7 S2 cells using the RNeasy Mini Kit (Qiagen). Gene expression analysis was performed using the Affymetrix (Santa Clara) *Drosophila* GeneChips according to the standard Affymetrix GeneChip protocol as outlined in the GeneChip Expression Analysis Technical Manual by Affymetrix (2001).

Luciferase reporter assay

Attacin reporter activity was measured essentially as described^{25,26}. We purchased *E. coli* from Molecular Probes and LPS from Sigma (St Louis). Of note, *Drosophila* is relatively resistant to LPS, as high concentrations are required to evoke biological responses²⁹.

Fly stocks

Fly culture and crosses were grown on standard fly medium at 25 °C. Oregon R were used as wild-type standard. The BG00650 line²⁸ was obtained from the Bloomington stock centre. *Dredd*^{B118}, *Relish*^{E20}, *Dif*¹ and *imd*¹ were described elsewhere^{29,30}. The P-element inserted in BG00650 was mobilized by following standard crosses. Mobilization events were screened by PCR using different primer combinations either in the genomic DNA flanking the P-element (5'-GACTGCCTGAAAATCGGACTC-3' and 5'-AGCACGCTT TTGAGATAACACC-3') or inside the P-element sequence (5'-CTTGCCGACGGGA CCACC-3'). $\Delta 5$ resulted from the imprecise excision of the P-element and deletes at least the first exon of *PGRP-LC*. *N18* contains a P-element insertion in the first intron of *PGRP-LC*.

Survival experiments

We performed bacterial challenge by pricking adult flies in the thorax with a sharpened tungsten needle dipped into a concentrated bacterial pellet (A_{150}). The time course of survival was carried out at 25 °C on a group of 15 flies from each line, and was repeated twice. We excluded flies that died after the first two hours of challenge from the analysis. Gram-negative bacterial *E. coli* strain 1106 and the Gram-positive bacterium *M. luteus* were used.

Northern blot and statistical analyses

We collected flies 10–12 h after infection and kept them at –80 °C until analysed. Total RNA extraction and northern blot were performed as described³¹. Data were analysed using one-way analysis of variance (ANOVA). A *P*-value of <0.05 was considered to be significant.

Received 1 March; accepted 7 March 2002.

Published online 24 March 2002, DOI 10.1038/nature735.

- Rizki, T. M. & Rizki, R. M. Properties of the larval hemocytes of *Drosophila melanogaster*. *Experimentia* **36**, 1223–1226 (1980).
- Franc, N. C., Heitzler, P., Ezekowitz, R. A. & White, K. Requirement for Croquemort in phagocytosis of apoptotic cells in *Drosophila*. *Science* **284**, 1991–1994 (1999).
- Aderem, A. & Underhill, D. M. Mechanisms of phagocytosis in macrophages. *Annu. Rev. Immunol.* **17**, 593–623 (1999).
- Garin, J. *et al.* The phagosome proteome: insight into phagosome functions. *J. Cell. Biol.* **152**, 165–180 (2001).
- Rämet, M. *et al.* *Drosophila* scavenger receptor Cl is a pattern recognition receptor for bacteria. *Immunity* **15**, 1027–1038 (2001).
- Hammond, S. M., Bernstein, E., Beach, D. & Hannon, G. J. An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. *Nature* **404**, 293–296 (2000).
- Pearson, A., Lux, A. & Krieger, M. Expression cloning of dSR-CI, a class C macrophage-specific scavenger receptor from *Drosophila melanogaster*. *Proc. Natl Acad. Sci. USA* **92**, 4056–4060 (1995).
- Bjerknes, R. Flow cytometric assay for combined measurement of phagocytosis and intracellular killing of *Candida albicans*. *J. Immunol. Methods* **72**, 229–241 (1984).
- Kirchhausen, T. Three ways to make a vesicle. *Nature Rev. Mol. Cell. Biol.* **1**, 187–198 (2000).
- Hackam, D. J. *et al.* Indirect role for COPI in the completion of Fcγ receptor-mediated phagocytosis. *J. Biol. Chem.* **276**, 18200–18208 (2001).
- Perry, D. G., Daugherty, G. L. & Martin, W. J. II Clathrin-coated pit-associated proteins are required for alveolar macrophage phagocytosis. *J. Immunol.* **162**, 380–386 (1999).
- Lebestky, T., Chang, T., Hartenstein, V. & Banerjee, U. Specification of *Drosophila* hematopoietic lineage by conserved transcription factors. *Science* **288**, 146–149 (2000).
- Rehorn, K. P., Thelen, H., Michelson, A. M. & Reuter, R. A molecular aspect of hematopoiesis and endoderm development common to vertebrates and *Drosophila*. *Development* **122**, 4023–4031 (1996).
- Oike, Y. *et al.* Mice homozygous for a truncated form of CREB-binding protein exhibit defects in hematopoiesis and vasculo-angiogenesis. *Blood* **93**, 2771–2779 (1999).
- Fehon, R. G., Dawson, I. A. & Artavanis-Tsakonas, S. A *Drosophila* homologue of membrane-skeleton protein 4.1 is associated with septate junctions and is encoded by the coracle gene. *Development* **120**, 545–557 (1994).
- Lee, S., Harris, K. L., Whittington, P. M. & Kolodziej, P. A *short stop* is allelic to *kakapo*, and encodes rod-like cytoskeletal-associated proteins required for axon extension. *J. Neurosci.* **20**, 1096–1108 (2000).
- Madden, K., Crowner, D. & Giniger, E. LOLA has the properties of a master regulator of axon-target interaction for SNB motor axons of *Drosophila*. *Dev. Biol.* **213**, 301–313 (1999).
- Juang, J. L. & Hoffmann, F. M. *Drosophila* Abelson interacting protein (dAbi) is a positive regulator of Abelson tyrosine kinase activity. *Oncogene* **18**, 5138–5147 (1999).
- Lanzetti, L. *et al.* The Eps8 protein coordinates EGF receptor signalling through Rac and trafficking through Rab5. *Nature* **408**, 374–377 (2000).
- Werner, T. *et al.* A family of Peptidoglycan recognition proteins in the fruit fly *Drosophila melanogaster*. *Proc. Natl Acad. Sci. USA* **97**, 13772–13777 (2000).
- Kang, D., Liu, G., Lundstrom, A., Gelius, E. & Steiner, H. A Peptidoglycan recognition protein in innate immunity conserved from insects to humans. *Proc. Natl Acad. Sci. USA* **95**, 10078–10082 (1998).

(1998).

- Ochiai, M. & Ashida, M. A pattern recognition protein for peptidoglycan. Cloning the cDNA and the gene of the silkworm, *Bombyx mori*. *J. Biol. Chem.* **274**, 11854–11858 (1999).
- Liu, C., Gelius, E., Liu, G., Steiner, H. & Dziarski, R. Mammalian Peptidoglycan recognition protein binds peptidoglycan with high affinity, is expressed in neutrophils, and inhibits bacterial growth. *J. Biol. Chem.* **275**, 24490–24499 (2000).
- Liu, C., Xu, Z., Gupta, D. & Dziarski, R. Peptidoglycan recognition proteins: a novel family of four human innate immunity pattern recognition molecules. *J. Biol. Chem.* **276**, 34686–34694 (2001).
- Michel, T., Reichhart, J. M., Hoffmann, J. A. & Royet, J. *Drosophila* Toll is activated by Gram-positive bacteria through a circulating Peptidoglycan recognition protein. *Nature* **414**, 756–759 (2001).
- Tauszig, S., Jouanguy, E., Hoffmann, J. A. & Imler, J. L. Toll-related receptors and the control of antimicrobial peptide expression in *Drosophila*. *Proc. Natl Acad. Sci. USA* **97**, 10520–10525 (2000).
- Imler, J. L., Tauszig, S., Jouanguy, E., Forestier, C. & Hoffmann, J. A. LPS-induced immune response in *Drosophila*. *J. Endotoxin Res.* **6**, 459–462 (2000).
- Lukacsovich, T. *et al.* Dual-tagging gene trap of novel genes in *Drosophila melanogaster*. *Genetics* **157**, 727–742 (2001).
- Imler, J. L. & Hoffmann, J. A. Signaling mechanisms in the antibacterial host defense of *Drosophila*. *Curr. Opin. Microbiol.* **3**, 16–22 (2000).
- Leulier, F., Rodriguez, A., Khush, R. S., Abrams, J. M. & Lemaitre, B. The *Drosophila* caspase Dredd is required to resist gram-negative bacterial infection. *EMBO Rep.* **1**, 353–358 (2000).
- Manfrulli, P., Reichhart, J. M., Steward, R., Hoffmann, J. A. & Lemaitre, B. A mosaic analysis in *Drosophila* fat body cells of the control of antimicrobial peptide genes by the Rel proteins Dorsal and DIF. *EMBO J.* **18**, 3380–3391 (1999).

Supplementary Information accompanies the paper on Nature's website (<http://www.nature.com>).

Acknowledgements

M.R. devised and carried out the RNAi screen and performed microarray and luciferase analysis. P.M. did the *in vivo* fly work under supervision of B.M.-P. A.P. did some of the targeted RNAi treatments and was intellectually involved throughout the project. M.R. and R.A.B.E. co-wrote the paper. We thank members of Hoffmann and Ezekowitz laboratories for discussions. We also thank M. Sackal for help with northern blots, J. Couget and M. Tahiliani for help with Affymetrix microarrays. M. Krieger provided the S2 cell cDNA library. This work was supported by grants from the National Institutes of Health (R.A.B.E.), the Foundation for Pediatric Research (M.R.), the Finnish Medical Foundation (M.R.), Maud Kuistila Foundation (M.R.), I.A.R.C. (P.M.) and American Cancer Society (A.P.).

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to MR (e-mail: mrarnet@partners.org).

Deubiquitination of p53 by HAUSP is an important pathway for p53 stabilization

Muyang Li*, Delin Chen*, Ariel Shiloh*, Jianyuan Luo*, Anatoly Y. Nikolaev*, Jun Qin† & Wei Gu*

* Institute for Cancer Genetics, and Department of Pathology, College of Physicians & Surgeons, Columbia University, 1150 St Nicholas Avenue, New York, New York 10032, USA

† Departments of Biochemistry and Cell Biology, Baylor College of Medicine, One Baylor Plaza, Houston, Texas 77030, USA

The p53 tumour suppressor is a short-lived protein that is maintained at low levels in normal cells by Mdm2-mediated ubiquitination and subsequent proteolysis^{1–3}. Stabilization of p53 is crucial for its tumour suppressor function^{1–5}. However, the precise mechanism by which ubiquitinated p53 levels are regulated *in vivo* is not completely understood. By mass spectrometry of affinity-purified p53-associated factors, we have identified herpesvirus-associated ubiquitin-specific protease⁶ (HAUSP) as a novel p53-interacting protein. HAUSP strongly stabilizes p53

even in the presence of excess Mdm2, and also induces p53-dependent cell growth repression and apoptosis. Significantly, HAUSP has an intrinsic enzymatic activity that specifically deubiquitinates p53 both *in vitro* and *in vivo*. In contrast, expression of a catalytically inactive point mutant of HAUSP in cells increases the levels of p53 ubiquitination and destabilizes p53. These findings reveal an important mechanism by which p53 can be stabilized by direct deubiquitination and also imply that HAUSP might function as a tumour suppressor *in vivo* through the stabilization of p53.

By serving as a signal for specific cellular protein degradation, ubiquitination is crucial in the physiological regulation of many cellular processes^{7–9}. The ubiquitination of p53 was first discovered in papilloma-virus-infected cells through the functions mediated by the viral E6 protein¹⁰, however, in normal cells, Mdm2 functions as a ubiquitin ligase (E3) that directly mediates the ubiquitination and subsequent degradation of p53 (refs 11–13). Numerous studies imply the existence of multiple pathways involved in p53 stabilization^{14–19}. In response to DNA damage, p53 is phosphorylated at multiple sites, and these phosphorylation events promote p53 stabilization by preventing the binding with Mdm2 and rendering p53 more resistant to Mdm2-mediated degradation^{14,15}. Furthermore, through inhibiting Mdm2-mediated ubiquitin ligase activity, the p14^{arf} tumour suppressor can stabilize p53 *in vivo* in response to

oncogene activation²⁰. Overall, regulation of the p53 ubiquitination levels is of intense interest but remains less well understood.

Using a biochemical purification method with glutathione *S*-transferase (GST)–p53 affinity chromatography (see Methods and refs 21,22), we have identified a novel p53-binding protein from nuclear extracts of human lung carcinoma cells (H1299). As indicated in Fig. 1a, there are several proteins present in the fractions eluted from the GST–p53 affinity column as well as other columns. Strikingly, only one protein, p135 (relative molecular mass ~135,000 (M_r 135K)) was specifically present in the associated factors obtained from the GST–p53 column but not from either the GST column or the control column (Fig. 1a, compare lane 3 with lanes 1 and 2). After a large preparation,

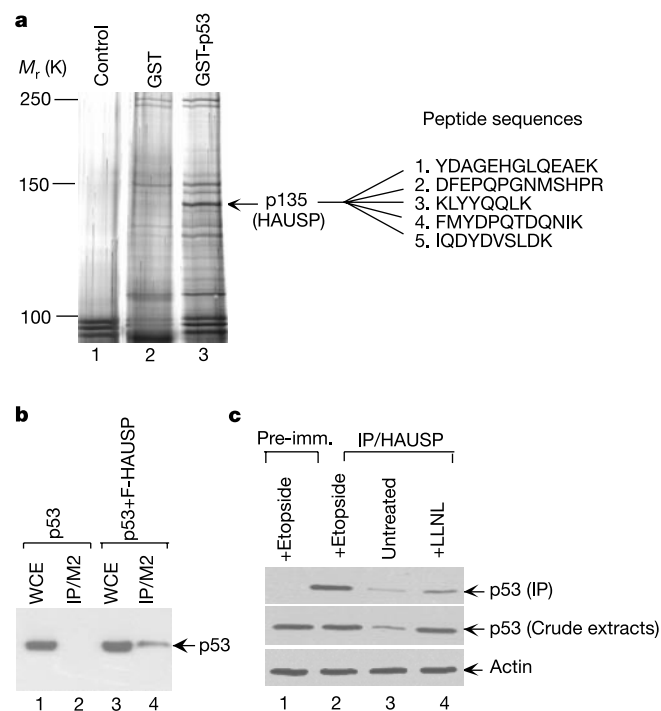


Figure 1 Purification of HAUSP, and interactions between p53 and HAUSP. **a**, Identification of HAUSP as a novel p53-binding protein. Silver staining analysis of an SDS–PAGE gel containing the eluates from the indicated columns. Peptide sequences derived from the p135 protein band were obtained by mass spectrometry. **b**, p53 interacts with HAUSP in cells. Western blot analysis of the whole cell extract (WCE) or immunoprecipitates (IP/M2) from the transfected cells by anti-p53 monoclonal antibody (DO-1). **c**, Interaction between endogenous p53 and HAUSP proteins. The top panel shows a western blot analysis of control immunoprecipitates with the pre-immune serum (lane 1) or immunoprecipitates with the anti-HAUSP antibody (IP/HAUSP) from H460 cells that were either untreated (lane 3) or treated with a DNA damage reagent (etoposide) (lane 2) or a proteasome inhibitor (LLNL) (lane 4). The middle and bottom panels show similar analyses of nuclear extracts by anti-p53 (middle) or anti-actin monoclonal antibody (bottom).

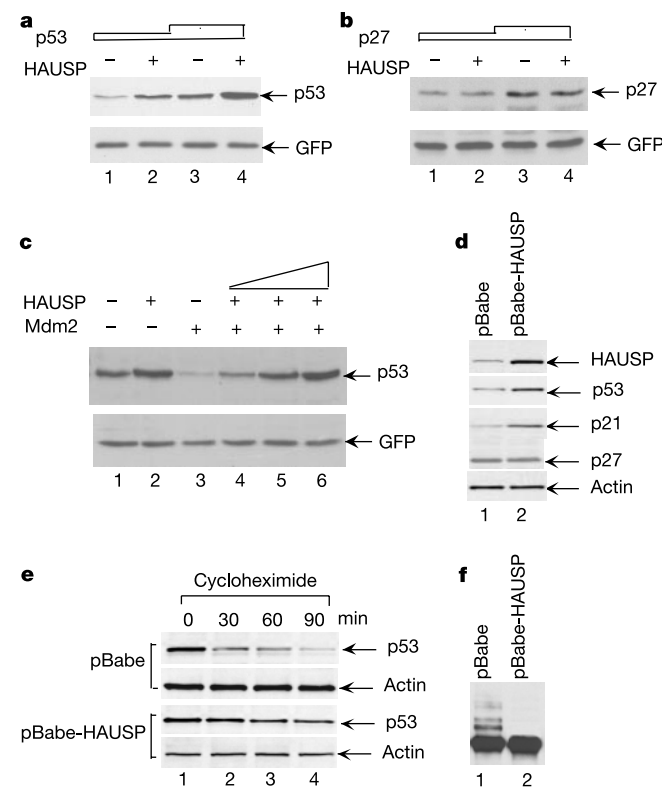


Figure 2 HAUSP interacts with and stabilizes p53 *in vivo*. **a**, **b**, Enhancement of the steady-state levels of p53 (**a**), but not p27 (**b**), by HAUSP. Western blot analysis of cell extracts from the H1299 cells transfected with p53 alone (lanes 1 and 3), or with p53 and HAUSP together (lanes 2 and 4), with anti-p53 monoclonal antibody (DO-1) (**a**). Western blot analysis of cell extracts from the H1299 cells transfected with p27 alone (lanes 1 and 3), or with p27 and HAUSP together (lanes 2 and 4), with anti-p27 monoclonal antibody (**b**). **c**, Protection of p53 from Mdm2-mediated degradation by HAUSP. Western blot analysis of extracts from the cells transfected with p53 (lane 1), with p53 and HAUSP together (lane 2) or with p53 and Mdm2 together (lane 3), or in combination with different amounts of HAUSP (lanes 4–6), with anti-p53 monoclonal antibody (DO-1). **d**, Regulation of the expression levels of endogenous proteins by HAUSP. Cell extracts from both mock-infected and pBabe-HAUSP-infected IMR-90 cells were analysed for expression levels of each protein by western blot analysis. **e**, Regulation of the half-life of endogenous p53 by HAUSP. Cell extracts from both mock-infected and pBabe-HAUSP-infected IMR-90 cells, harvested at different time points as indicated after pretreatment with cycloheximide were analysed for p53 protein levels by western blotting with anti-p53 monoclonal antibody (DO-1). **f**, Regulation of the ubiquitination levels of endogenous p53 by HAUSP. Cell extracts from both mock-infected cells and pBabe-HAUSP-infected IMR-90 cells pretreated with LLNL for 4 h were first immunoprecipitated with anti-p53 polyclonal antibody, then analysed for ubiquitination levels by western blotting with anti-p53 monoclonal antibody (DO-1).

enough material of the p135 band was obtained for mass spectrometry; five peptide sequences were obtained, all of which were derived from the herpesvirus protein Vmw110-associated cellular factor known as HAUSP (also known as human USP7 (ref. 6)). HAUSP belongs to the ubiquitin-specific processing protease (UBP) family of deubiquitination enzymes (DUBs) and contains the characteristic Cys and His motifs at the core enzymatic domain⁶. Interestingly, the amino-terminal and carboxy-terminal extensions of HAUSP with no significant homology to other members of the UBP family, which are thought to be critical for the substrate specificity^{6,23–25}, bind directly to p53 *in vitro* (see Supplementary Information). To evaluate interactions *in vivo* by immunoprecipitation analysis, p53-null cells (H1299) were transfected with p53 and a Flag-tagged HAUSP expression vector. As shown in Fig. 1b, p53 was readily immunoprecipitated from the cells transfected with both Flag-HAUSP and p53 (lane 4) but not from cells transfected with p53 alone (lane 2). By using the HAUSP-specific antibody, we also examined the interaction between the endogenous p53 and HAUSP proteins. Western blot analysis showed that p53 was present in the anti-HAUSP immunoprecipitates from cell extracts of human lung carcinoma cells (H460), but not in the control immunoprecipitates obtained with the preimmune serum (Fig. 1c). This interaction was strongly detected in cells subjected to genotoxic stress (Fig. 1c, compare lanes 2 and 3), whereas only a slight enhancement was detected in the cells treated with a proteasome inhibitor LLNL (lane 4). These results indicate that p53 interacts with HAUSP *in vivo* and that the possible regulation of p53 by HAUSP might be still effective in the cells during the DNA damage response.

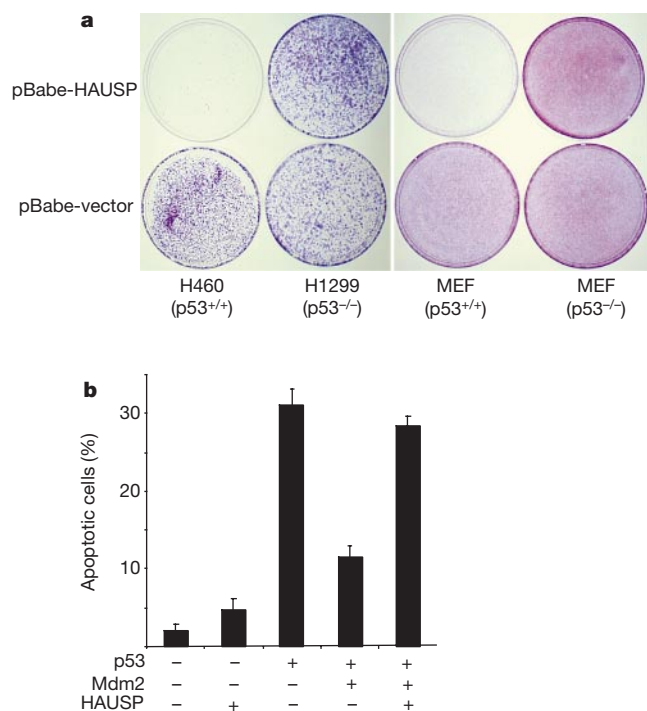


Figure 3 Effects of HAUSP on p53-mediated cell growth repression (**a**) and apoptosis (**b**). **a**, A pair of human lung carcinoma cells (H1299 (p53^{+/+}) and H460 (p53^{-/-})) and a pair of mouse embryo fibroblasts (MEF (p53^{+/+}) and MEF (p53^{-/-})) were infected with either pBabe-vector or pBabe-HAUSP. At 24 h after infection, cells were split and kept in the medium with puromycin, and surviving colonies were counted after 2 weeks. **b**, H1299 cells were transfected with p53 alone, with HAUSP alone, with p53 and Mdm2 together or with p53, Mdm2 and HAUSP together, as indicated. After transfection the cells were fixed, stained for p53 with fluorescein-isothiocyanate-conjugated anti-p53 antibody and analysed for apoptotic cells (sub-G1) according to DNA content (propidium iodide staining).

To determine the functional consequence of the p53–HAUSP interaction, we tested whether HAUSP affects stabilization of p53. As indicated in Fig. 2a, HAUSP expression significantly increased the steady-state cellular levels of p53. In contrast, HAUSP had no obvious effect on the levels of p27 (Fig. 2b), another short-lived tumour suppressor protein whose stability is also regulated by the ubiquitination pathway²⁶. Moreover, as shown in Fig. 2c, HAUSP effectively rescues p53 from Mdm2-mediated degradation. Thus, although overexpression of Mdm2 significantly induced p53 degradation (compare lane 3 with lane 1), degradation of p53 was inhibited in a dose-dependent manner upon expression of HAUSP (lanes 4–6). Furthermore, we examined the effect of HAUSP expression on stabilization of endogenous p53. Normal human fibroblast IMR-90 cells were infected with either a pBabe retrovirus empty vector or a pBabe retrovirus containing HAUSP. We first examined the protein levels of endogenous p53 by western

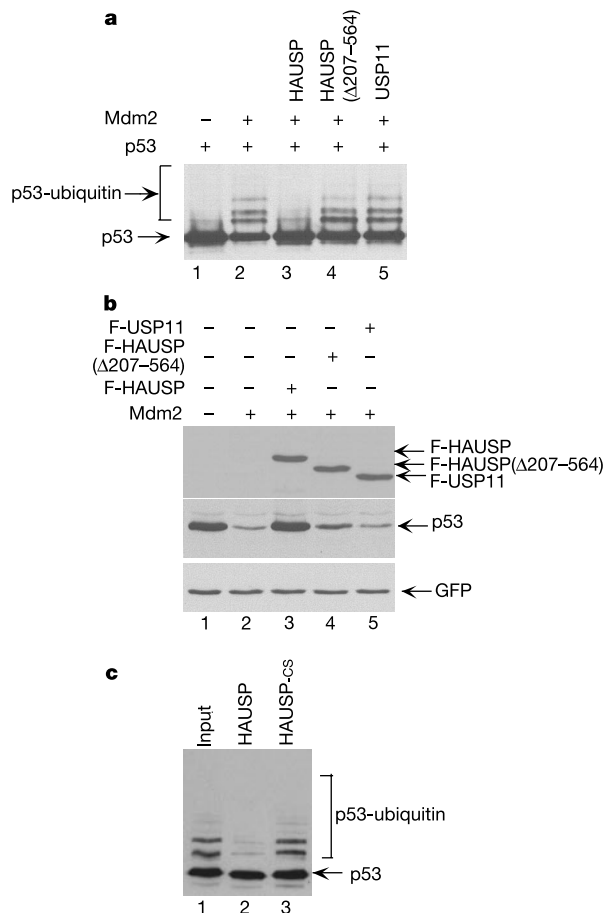


Figure 4 Deubiquitination of p53 by HAUSP both *in vivo* and *in vitro*. **a**, Regulation of p53 ubiquitination levels *in vivo*. Western blot analysis of immunoprecipitates with the M2/Flag antibody from the cells transfected with Flag-p53 (lane 1), with Flag-p53 and Mdm2 together (lane 2), or in combination with different expression vectors as indicated (lanes 3–5), with anti-p53 monoclonal antibody (DO-1). **b**, Regulation of p53 stability by the HAUSP mutant. Western blot analysis of H1299 cell extracts from the cells transfected with p53 (lane 1), with p53 and Mdm2 together (lane 2), or in combination with different expression vectors as indicated (lanes 3–5), with anti-p53 monoclonal antibody (DO-1). The CMV–GFP expression vector was included in each transfection as a transfection efficiency control, and levels of GFP were detected with anti-GFP monoclonal antibody (JL-8; Clontech). **c**, Deubiquitination of p53 *in vitro* by HAUSP. The purified ubiquitinated p53 protein (lane 1) was incubated with the purified recombinant proteins of either HAUSP (lane 2) or HAUSP-cs (lane 3).

blot analysis. Significantly higher levels of p53 proteins were detected in the pBabe-HAUSP-infected cells (Fig. 2d, compare lanes 2 and 1). Interestingly, expression of endogenous p21 was also induced (compare lanes 2 and 1), which is consistent with transient transfection results (see Supplementary Information), indicating that HAUSP also activates p53-dependent transcriptional activation. In contrast, the levels of endogenous p27 remained the same, supporting the notion that HAUSP stabilizes p53 but not p27 *in vivo*. Notably, the half-life of p53 in the pBabe-HAUSP-infected cells was significantly increased by HAUSP expression (~90 min), whereas the half-life of p53 in the mock-infected cells was less than 30 min (Fig. 2e). In corroboration of these results, we found that the ubiquitination levels of p53 in the pBabe-HAUSP-infected cells were also reduced in comparison with the levels in the mock-infected cells (Fig. 2f). These data therefore demonstrate that HAUSP specifically stabilizes p53 *in vivo*.

To investigate the biological role of HAUSP, we examined its effect on cell growth in a colony formation assay. A pair of human lung carcinoma cells (H1299 and H460) were infected with either an empty pBabe-puro control retrovirus or a pBabe-puro retrovirus encoding HAUSP, and cultured for 2 weeks under pharmacological selection. Strikingly, HAUSP strongly inhibited the growth of H460 cells expressing wild-type p53 but had no significant effect on p53-null H1299 cells (Fig. 3a). Similar cell growth repression by HAUSP

was also observed in MEF p53^{+/+} cells but not MEF p53^{-/-} cells (Fig. 3a), indicating that the cell growth repression by HAUSP is dependent on p53. We also tested whether HAUSP directly affects p53-dependent apoptosis. H1299 cells were transfected with p53 alone, with p53 and Mdm2, or with p53, Mdm2 and HAUSP. After transfection the cells were fixed, stained for p53, and analysed for apoptotic cells (sub-G1) according to DNA content²². As indicated in Fig. 3b, although overexpression of p53 alone induced significant apoptosis (31.0%), Mdm2 strongly reduced the level of p53-dependent apoptosis (11.2%). However, expression of HAUSP effectively attenuated the inhibitory effect of Mdm2 on p53-mediated apoptosis (28.5% versus 11.2%; Fig. 3b). These data demonstrate that HAUSP is crucially involved in both the regulation of p53-dependent apoptosis and the inhibition of cell growth.

To elucidate the molecular mechanism by which HAUSP stabilizes p53, we tested whether HAUSP directly controls the levels of p53 ubiquitination *in vivo*. As indicated in Fig. 4a, a high level of ubiquitinated p53 was found in cells transfected with p53 and Mdm2 (lane 2); however, p53 ubiquitination was significantly abrogated by HAUSP expression (compare lanes 3 and 2). In contrast, HAUSP had no effect on the levels of ubiquitinated p27 (see Supplementary Information). Notably, an unrelated human UB family member (human USP11) (refs 23–25) that is defective in p53 binding (see Supplementary Information) had no obvious effect on the levels of p53 ubiquitination (Fig. 4a, lane 5) or stabilizing p53 (Fig. 4b, lane 5). Significantly, a HAUSP mutant with a short deletion at the core domain lost the ability both to stabilize p53 (Fig. 4b, lane 4) and to reduce the cellular levels of p53 ubiquitination (Fig. 4a, lane 4), indicating that stabilization of p53 by HAUSP requires its deubiquitinating enzymatic activity. To confirm the specific deubiquitination activity of HAUSP on p53, we examined whether HAUSP can directly deubiquitinate p53 in a purified system. The HAUSP protein was expressed in bacteria and purified to near homogeneity. The ubiquitinated form of p53 was purified on an M2 affinity column under conditions of high stringency from cells transfected with a Flag-tagged p53 expression vector. The highly purified *in vitro* system was used in this assay to avoid possible contamination by either inhibitory factors (namely, p14^{arr}) or any enzymes involving the ubiquitination of p53. As shown in Fig. 4c, p53 was efficiently deubiquitinated upon incubation with purified recombinant HAUSP (lane 2). These results therefore demonstrate that HAUSP can specifically deubiquitinate p53 both *in vitro* and *in vivo*.

Interestingly, HAUSP-cs, a point mutant of HAUSP in which a highly conserved Cys residue at the core domain was replaced by Ser, retained its strong binding with p53 (see Supplementary Information); however, it was functionally defective in deubiquitinating p53 *in vitro* (Fig. 4c, lane 3). Significantly, in contrast to the effect of wild-type HAUSP, the expression of HAUSP-cs in the cells increased the level of p53 ubiquitination (Fig. 5a, compare lanes 4 and 2), indicating that HAUSP-cs might function as a dominant-negative mutant through interfering with endogenous HAUSP-mediated deubiquitination of p53. To corroborate these results, we also tested whether HAUSP-cs expression affects the levels of p53 proteins in cells. As shown in Fig. 5b, expression of HAUSP-cs together with p53 slightly, but significantly, decreases the levels of p53 proteins (compare lanes 1 and 3 with lanes 2 and 4). To demonstrate further that HAUSP regulates endogenous p53, we introduced HAUSP-cs into normal human cells: IMR-90 cells were infected with either a pBabe retrovirus empty vector or a pBabe retrovirus containing HAUSP-cs. As indicated in Fig. 5c, in the mock-infected cells the level of p53 proteins was markedly increased by DNA damage (lanes 1, 3 and 5); however, HAUSP-cs expression led to a significant attenuation of p53 stabilization under both normal and DNA-damage conditions (lanes 2, 4 and 6). Taken together, these results indicate that HAUSP is crucially involved in both the deubiquitination and the stabilization of p53 under physiological conditions.

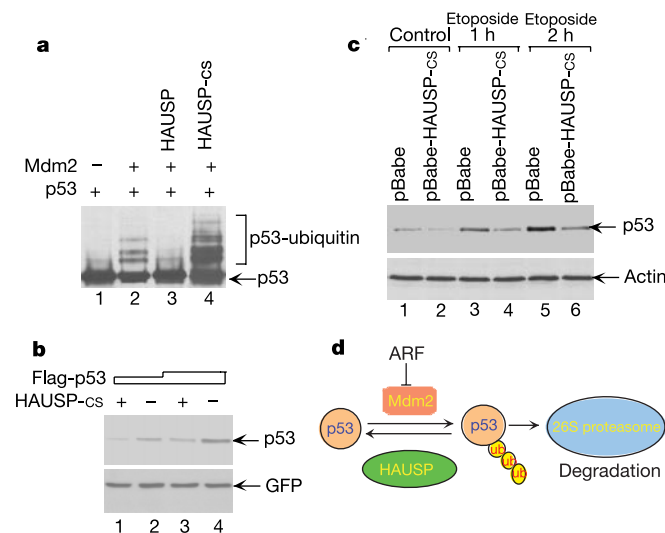


Figure 5 The dominant-negative effects of HAUSP-cs in human cells. **a**, Western blot analysis of immunoprecipitates with the M2/Flag antibody from the cells transfected with Flag-p53 (lane 1), with Flag-p53 and Mdm2 together (lane 2), or in combination with HAUSP and HAUSP-cs as indicated (lanes 3 and 4), with anti-p53 monoclonal antibody (DO-1). All cells were treated with 50 μ M LLNL for 4 h before being harvested. **b**, Western blot analysis of immunoprecipitates with M2/Flag antibody from human JSJA cells either transfected with expression vectors of Flag-p53 alone (lanes 2 and 4) or with expression vectors of Flag-p53 and HAUSP-cs together (lanes 1 and 3), with anti-p53 monoclonal antibody (DO-1). The CMV-GFP expression vector was included in each transfection as a transfection efficiency control, and levels of GFP were detected with anti-GFP monoclonal antibody (JL-8; Clontech). **c**, Western blot analysis of the cell extracts from both mock-infected and pBabe-HAUSP-cs-infected IMR-90 cells with anti-p53 monoclonal antibody (DO-1). Cells were either not treated (lanes 1 and 2) or treated with 20 μ M etoposide (lanes 3–6) for either 1 or 2 h as indicated. **d**, A model for the regulation of p53 stability by Mdm2, HAUSP and ARF. p53 is ubiquitinated (ub) by Mdm2 and subsequently degraded by the 26S proteasome, whereas the ARF tumour suppressor induces p53 stabilization through inhibiting Mdm2-mediated ubiquitin ligase activity. HAUSP can deubiquitinate p53 directly and rescue the ubiquitinated p53 from degradation.

Our data suggest that HAUSP-mediated stabilization of p53 acts through its intrinsic deubiquitinating enzymatic activity. Deubiquitination, which removes the ubiquitin moiety from ubiquitin-modified proteins, is now recognized as an important regulatory step^{23–25}. The large number of UBP proteins also indicates that they might bind to specific cellular proteins and have substrate specificity^{23–25}. Although an increasing number of UBP homologues have been identified in mammalian cells^{23–25}, none has yet been implicated in stabilizing specific substrates *in vivo*. HAUSP might represent the first example of a mammalian protein that can directly deubiquitinate and stabilize a specific cellular factor (p53). Our findings have significant implications for the potential tumour suppression function of HAUSP and also predict that many UBP family proteins, like HAUSP, might interact with different substrates *in vivo* for deubiquitination as well as subsequent protein stabilization.

Previous studies have indicated that HAUSP interacts with the herpesvirus protein Vmw110 and that a subset of the HAUSP proteins is localized with the promyelocytic leukaemia (PML) nuclear body⁶. There is also now increasing evidence showing that ARF is dispensable for p53 activation induced by some types of oncogenic stress^{27,28}. These studies, together with our findings, therefore further indicate potential regulation of the p53–HAUSP interaction in viral infection, DNA damage response and other types of stress response. Stabilization of p53 is crucial for its effects on cell growth repression and apoptosis. Many studies have proposed that stabilization of p53 in response to various types of stress can be achieved through inhibition of the Mdm2–p53 interaction and/or Mdm2-mediated ubiquitin ligase^{1–5,20}. Our findings reveal that ubiquitination of p53 is a dynamic process *in vivo* and that ubiquitinated p53 can be rescued from degradation by HAUSP through direct deubiquitination (Fig. 5d). It is very likely that changing the balance between the Mdm2-mediated ubiquitination and HAUSP-mediated deubiquitination of p53 is the key for p53 stabilization *in vivo*. □

Methods

Plasmids and antibodies

To construct the HAUSP and USP11 expression vectors, the DNA sequences corresponding to the full-length proteins^{6,24} were amplified by polymerase chain reaction (PCR) from Marathon-Ready Hela cDNA (Clontech), and subcloned either into a pcDNA3-Topo vector (Invitrogen) or with a Flag-tag into either a pET-11 vector for expression in bacteria or a pCIN4 vector and pBabe vector for expression in mammalian cells^{22,29}. For the different deletion mutant constructs, DNA sequences corresponding to different regions were amplified by PCR from the above constructs and subcloned into respective expression vectors. To prepare the HAUSP antibody, we made a polyclonal antibody against the recombinant HAUSP full-length protein. The DNA sequence corresponding to the full-length protein was subcloned into the pET-15-His vector (Novagene). Anti-HAUSP antisera were raised in rabbits against the purified His-HAUSP protein (Covance), and further affinity-purified on a protein A column.

Identification of HAUSP as a p53-binding protein

Eliminating non-specific protein binding is critical for successfully identifying a genuine binding partner for p53 in the affinity chromatography assay^{21,22}. We modified the salt concentration range between the binding and elution conditions compared with the previous method^{21,22} (for example, 200 mM NaCl for binding, 500 mM NaCl for elution) to limit the number of proteins eluted. Furthermore, nuclear extracts were extensively precleared with the GST column before being loaded on the GST–p53 affinity column. The mock purification was performed simultaneously on both the GST column loaded with the same nuclear extract and another GST–p53 column loaded with a blank buffer for identifying any possible non-specific binding. In brief, a column containing 40 µl of the indicated GST fusion protein coupled to beads was washed extensively with BC500 buffer (25 mM Tris-HCl pH 7.8, 500 mM NaCl, 1 mM EDTA, 1 mM dithiothreitol (DTT), 10% glycerol, 0.2% Nonidet P40 (NP40), 1% Triton X-100, 0.1% sodium deoxycholate), and then equilibrated with BC200 (25 mM Tris-HCl pH 7.8, 200 mM NaCl, 1 mM EDTA, 1 mM DTT, 10% glycerol, 0.2% NP40) for loading. Human lung carcinoma cells (H1299) were expanded in DMEM medium, and nuclear extracts were prepared essentially as described previously^{21,22}. Nuclear extracts were adjusted to 200 mM NaCl and 0.2% NP40 and precleared by flowing through the GST column for at least three times. Precleared nuclear extract (800 µl) was loaded on either the GST or the GST–p53 mini-column. After being washed five times with 1 ml BC200 buffer, the associated proteins were eluted from

the column with 40 µl BC500. The nuclear extract derived from $\sim 10 \times 10^9$ cells was used for the large preparation.

Stabilization of p53 and detection of ubiquitination levels of p53 *in vivo* and *in vitro*

The p53-null H1299 cells were transfected with 0.1–2 µg cytomegalovirus (CMV)–Flag–p53, 2 µg CMV–Mdm2, 1 µg CMV–green fluorescent protein (CMV–GFP), and 5–16 µg either CMV–HAUSP or the same quantities of expression vectors for the indicated HAUSP mutants or other proteins. At 24 h after transfection, cells were lysed in a Flag-lysis buffer (50 mM Tris-HCl pH 7.8, 137 mM NaCl, 10 mM NaF, 1 mM EDTA, 1% Triton X-100, 0.2% Sarkosyl, 1 mM DTT, 10% glycerol and fresh proteinase inhibitors) for western blot analysis. The levels of GFP were detected with the anti-GFP monoclonal antibody (JL-8; Clontech) as a transfection efficiency control. The ubiquitination levels of p53 were detected essentially as described previously³⁰. The cells were treated for 4 h with a proteasome inhibitor, LLNL (Sigma) (50 µM) before being harvested, and were then lysed in Flag-lysis buffer with mild sonication. The cell extracts were immunoprecipitated with Flag monoclonal antibody (M2), and subsequently resolved by SDS–polyacrylamide gel electrophoresis (8% or 4–20% gel) (Novex) and analysed by western blotting with anti-p53 (DO-1). For the preparation of a large amount of ubiquitinated p53 as the substrate for the deubiquitination assay *in vitro*, H1299 cells (5×10^7) were transfected together with the Flag–p53 and Mdm2 expression vectors. After treatment as described above, the ubiquitinated p53 was purified from the cell extracts on the M2-affinity column with the Flag-lysis buffer. After extensive washing with the Flag-lysis buffer, the proteins were eluted in BC100 buffer (25 mM Tris-HCl pH 7.8, 100 mM NaCl) with Flag-peptides (Sigma). The recombinant Flag–HAUSP and the mutant form, HAUSP–cs, were expressed in BL21 cells and purified on the M2 column. For the deubiquitination assay reaction *in vitro*, the ubiquitinated p53 protein was incubated with recombinant HAUSP (100 ng) or the same amount of other indicated proteins in a deubiquitination buffer (50 mM Tris-HCl pH 8.0, 50 mM NaCl, 1 mM EDTA, 10 mM DTT, 5% glycerol) for 2 h at 37 °C.

Received 17 December 2001; accepted 22 February 2002.

Published online 31 March 2002, DOI 10.1038/nature737.

- Ashcroft, M. & Vousden, K. H. Regulation of p53 stability. *Oncogene* **18**, 7637–7643 (1999).
- Oren, M. Regulation of the p53 tumour suppressor protein. *J. Biol. Chem.* **274**, 36031–36034 (1999).
- Freedman, D. A., Wu, L. & Levine, A. J. Functions of the MDM2 oncoprotein. *Cell Mol. Life Sci.* **55**, 96–107 (1999).
- Prives, C. & Hall, P. A. The p53 pathway. *J. Pathol.* **187**, 112–126 (1999).
- Vogelstein, B., Lane, D. & Levine, A. J. Surfing the p53 network. *Nature* **408**, 307–310 (2000).
- Everett, R. D. et al. A novel ubiquitin-specific protease is dynamically associated with the PML nuclear domain and binds to a herpesvirus regulatory protein. *EMBO J.* **16**, 566–577 (1997).
- Hershko, A., Ciechanover, A. & Varshavsky, A. The ubiquitin system. *Nature Med.* **6**, 1073–1081 (2000).
- Laney, J. D. & Hochstrasser, M. Substrate targeting in the ubiquitin system. *Cell* **97**, 427–430 (1999).
- Kornitzer, D. & Ciechanover, A. Modes of regulation of ubiquitin-mediated protein degradation. *J. Cell. Physiol.* **182**, 1–11 (2000).
- Scheffner, M., Huibregt, J. M., Vierstra, R. D. & Howley, P. The HPV-16 E6 and E6-AP complex functions as a ubiquitin-protein ligase in the ubiquitination of p53. *Cell* **75**, 495–505 (1993).
- Haupt, Y., Maya, R., Kazan, A. & Oren, M. Mdm2 promotes the rapid degradation of p53. *Nature* **387**, 296–299 (1997).
- Kubbutat, M. H. G., Jones, S. N. & Vousden, S. N. Regulation of p53 stability by Mdm2. *Nature* **387**, 299–303 (1997).
- Honda, R., Tanaka, H. & Yasuda, H. Oncoprotein MDM2 is a ubiquitin ligase E3 for tumor suppressor p53. *FEBS Lett.* **420**, 25–27 (1997).
- Shieh, S. Y., Ikeda, M., Taya, Y. & Prives, C. DNA damage-induced phosphorylation of p53 alleviates inhibition MDM2. *Cell* **91**, 325–334 (1997).
- Appella, E. & Anderson, C. W. Signaling to p53: breaking the posttranslational modification code. *Pathol. Biol. (Paris)* **48**, 227–245 (2000).
- Ashcroft, M., Kubbutat, M. H. & Vousden, K. H. Regulation of p53 function and stability by phosphorylation. *Mol. Cell. Biol.* **19**, 1751–1758 (1999).
- Blattner, C., Tobiasch, E., Litfen, M., Rahmsdorf, H. J. & Herrlich, P. DNA damage induced p53 stabilization: no indication for an involvement of p53 phosphorylation. *Oncogene* **18**, 1723–1732 (1999).
- Dumaz, N. & Meek, D. W. Serine15 phosphorylation stimulates p53 transactivation but does not directly influence interaction with HDM2. *EMBO J.* **18**, 7002–7010 (1999).
- Ashcroft, M., Taya, Y. & Vousden, K. H. Stress signals utilize multiple pathways to stabilize p53. *Mol. Cell. Biol.* **20**, 3224–3233 (2000).
- Sherr, C. J. & Weber, J. D. The ARF/p53 pathway. *Curr. Opin. Genet. Dev.* **10**, 94–99 (2000).
- Gu, W., Shi, X. L. & Roeder, R. G. Synergistic activation of transcription by CBP and p53. *Nature* **387**, 819–823 (1997).
- Luo, J., Su, F., Chen, D., Shiloh, A. & Gu, W. Deacetylation of p53 modulates its effect on cell growth and apoptosis. *Nature* **408**, 377–381 (2000).
- D'Andrea, A. & Pellmar, D. Deubiquitinating enzymes: a new class of biological regulators. *Crit. Rev. Biochem. Mol. Biol.* **35**, 337–352 (1998).
- Chung, C. H. & Baek, S. H. Deubiquitinating enzymes: their diversity and emerging roles. *Biochem. Biophys. Res. Commun.* **266**, 633–640 (1999).
- Wilkinson, K. D. Ubiquitination and deubiquitination: targeting of proteins for degradation by the proteasome. *Semin. Cell Dev. Biol.* **11**, 141–148 (2000).
- Slingerland, J. & Pagano, M. Regulation of the cdk inhibitor p27 and its deregulation in cancer. *J. Cell Physiol.* **183**, 10–17 (2000).
- Seavey, S. E., Holubar, M., Saucedo, L. J. & Perry, M. E. The E7 oncoprotein of human papillomavirus type 16 stabilizes p53 through a mechanism independent of p19 (ARF). *J. Virol.* **73**, 7590–7598 (1999).
- Tolbert, D., Lu, X., Yin, C., Tantama, M. & Van Dyke, T. p19ARF is dispensable for oncogenic stress-induced p53-mediated apoptosis and tumor suppression *in vivo*. *Mol. Cell. Biol.* **22**, 370–377 (2002).

29. Luo, J. *et al.* Negative control of p53 by Sir2 α promotes cell survival under stress. *Cell* **107**, 137–148 (2001).
30. Rodriguez, M. S., Desterro, J. M. P., Lain, S., Lane, D. P. & Hay, R. T. Multiple C-terminal lysine residues target p53 for ubiquitin-proteasome-mediated degradation. *Mol. Cell. Biol.* **20**, 8458–8467 (2000).

Supplementary Information accompanies the paper on Nature's website (<http://www.nature.com>).

Acknowledgements

We thank R. Baer, R. Dalla-Favera, B. Tycko and T. Ludwig for critical discussions; we also thank many colleges in the field for providing antibodies, cell lines and plasmids, and other members of W.G.'s laboratory for sharing unpublished data and critical comments. This work was supported in part by grants from Avon Foundation, the Stewart Trust, the Irma T. Hirsch Trust and NIH/NCI to W.G., who is also a Leukemia and Lymphoma Society Scholar.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for material should be addressed to W.G. (e-mail: wg8@columbia.edu).

Dissecting glucose signalling with diversity-oriented synthesis and small-molecule microarrays

Finny G. Kuruvilla*, Alykhan F. Shamji*[‡], Scott M. Sternson*[‡], Paul J. Hergenrother*[§] & Stuart L. Schreiber*

* Howard Hughes Medical Institute, Institute for Chemistry and Cell Biology, Bauer Center for Genomics Research, Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford Street, Cambridge, Massachusetts 02138, USA

[‡] Department of Biophysics, Harvard University, Cambridge, Massachusetts 02138, USA

Small molecules that alter protein function provide a means to modulate biological networks with temporal resolution. Here we demonstrate a potentially general and scalable method of identifying such molecules by application to a particular protein, Ure2p, which represses the transcription factors Gln3p and Nil1p^{1–3}. By probing a high-density microarray of small molecules generated by diversity-oriented synthesis with fluorescently labelled Ure2p, we performed 3,780 protein-binding assays in parallel and identified several compounds that bind Ure2p. One compound, which we call uretupamine, specifically activates a glucose-sensitive transcriptional pathway downstream of Ure2p. Whole-genome transcription profiling and chemical epistasis demonstrate the remarkable Ure2p specificity of uretupamine and its ability to modulate the glucose-sensitive subset of genes downstream of Ure2p. These results demonstrate that diversity-oriented synthesis and small-molecule microarrays can be used to identify small molecules that bind to a protein of interest, and that these small molecules can regulate specific functions of the protein.

The progress in identifying and expressing all human proteins⁴ presents an opportunity to develop a small-molecule modulator for every protein function. Small-molecule approaches to study protein function have illuminated diverse fields of biology. Examples

include tetrodotoxin, which enabled the dissection of the action potential⁵, and agonists of peroxisome-proliferator-activated receptor- γ such as rosiglitazone, which illuminated the regulation of adipogenesis⁶. However, in most cases no small molecule that can modulate the function of a protein of interest is known, and there is currently no efficient method of identifying these biological probes. Using the example of the yeast protein Ure2p, we demonstrate a general two-step method that does not require a high-resolution structure or a previously characterized small molecule known to bind the protein. First, diversity-oriented synthesis is used to produce structurally complex and diverse small molecules efficiently. Second, the resulting compounds are screened for their ability to bind a protein of interest by using small-molecule microarrays, a technique for extremely high-throughput parallel-binding assays. Cell-based studies can subsequently determine which functions of the protein are modulated by each small molecule.

The yeast protein Ure2p has been widely studied in several different contexts. Ure2p is the central repressor of genes involved

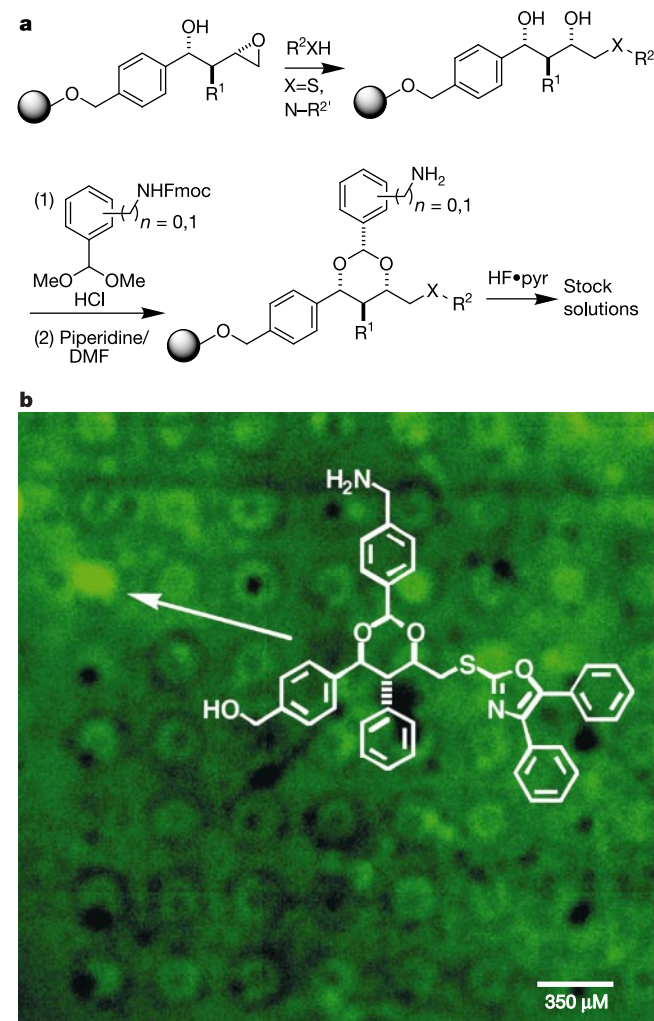


Figure 1 The library synthesis and identification of uretupamine. **a**, Outline of the diversity-oriented synthesis leading to uretupamine and other library members¹¹. **b**, An expanded view of 64 compound spots on the 3,780-member small-molecule microarray (~800 spots cm⁻²). Cy5-labelled Ure2p was passed over a microarray of the 1,3-dioxane small-molecule library, and the resulting slide was washed three times and scanned for fluorescence. The spot corresponding to uretupamine A is shown.

[‡] These authors contributed equally to this work.

[§] Present address: Department of Chemistry, University of Illinois, Urbana, Illinois 61801, USA.

in nitrogen metabolism⁷, is capable of switching to a prion form⁸, and is part of a signalling cascade downstream of the Tor proteins^{9,10}. Because there is no known small molecule that binds to Ure2p, we screened a collection of 3,780 structurally complex 1,3-dioxane small molecules resulting from a diversity-oriented synthesis^{11,12} (Fig. 1a). The molecules are structurally unbiased towards any particular protein target and can be used to identify specific probes for many different proteins. This collection of molecules had been prepared with a 'one bead-one stock solution' technology platform with the use of macrobeads (Fig. 1a), followed by automated compound cleavage and the generation of 5-mM stock solutions in 5 μ l of *N,N*-dimethylformamide^{11,13,14}. The small molecules were arrayed in high-density on glass slides ($\sim$ 800 spots cm^{-2} , 1 nl of each compound per slide) with a quill-pin contact-printing robot^{15,16}. These microarrays were probed with fluorescently labelled Ure2p, enabling the protein-binding properties of each molecule to be tested in parallel with minimal protein consumption (protein solution: 20 $\mu\text{g ml}^{-1}$, 0.2 ml). This method has been used to detect known interactions such as that between FKBP12 and a synthetic pipecolyl α -ketoamide^{15,16} and is applied here to the identification of novel small-molecule-protein interactions with uncharacterized compounds. Eight compound spots on the microarrays showed reproducible binding to labelled Ure2p (see Fig. 1b for one such spot in an 8 $\times$ 8 spot array where each spot

is derived from a single-compound stock solution derived from the diversity-oriented synthesis).

To determine cellular activity, the molecules comprising these spots were resynthesized and tested for the modulation of endogenous Ure2p function with a *PUT1-lacZ* reporter system because *PUT1* expression is known to be repressed by *URE2* (ref. 17). In addition to the positive control of rapamycin⁹, one of the eight compounds activated this reporter (Fig. 2a). The compound, which we named uretupamine A, gave a concentration-dependent dose response that at higher concentrations approached the levels of reporter gene activation induced by rapamycin (Fig. 2b).

To explore structure-activity relationships, we synthesized a series of compounds with systematic variations in the structure of uretupamine A (Fig. 2c). Specific atomic interactions are responsible for uretupamine binding because most modifications of its structure resulted in a complete loss of activity (Fig. 2c). Uretupamine A was rendered functionally inactive by acylation of the primary amine, replacement of the diphenyloxazole moiety with a phenyl group or a benzoxazole group or modification of the benzyl alcohol moiety (Fig. 2c). However, the C-5 position of the dioxane ring was tolerant to modification. Because the potency of uretupamine A was attenuated by poor solubility at higher concentrations in yeast medium, we synthesized a more soluble derivative, uretupamine B, which lacked the C-5 phenyl group on the dioxane ring

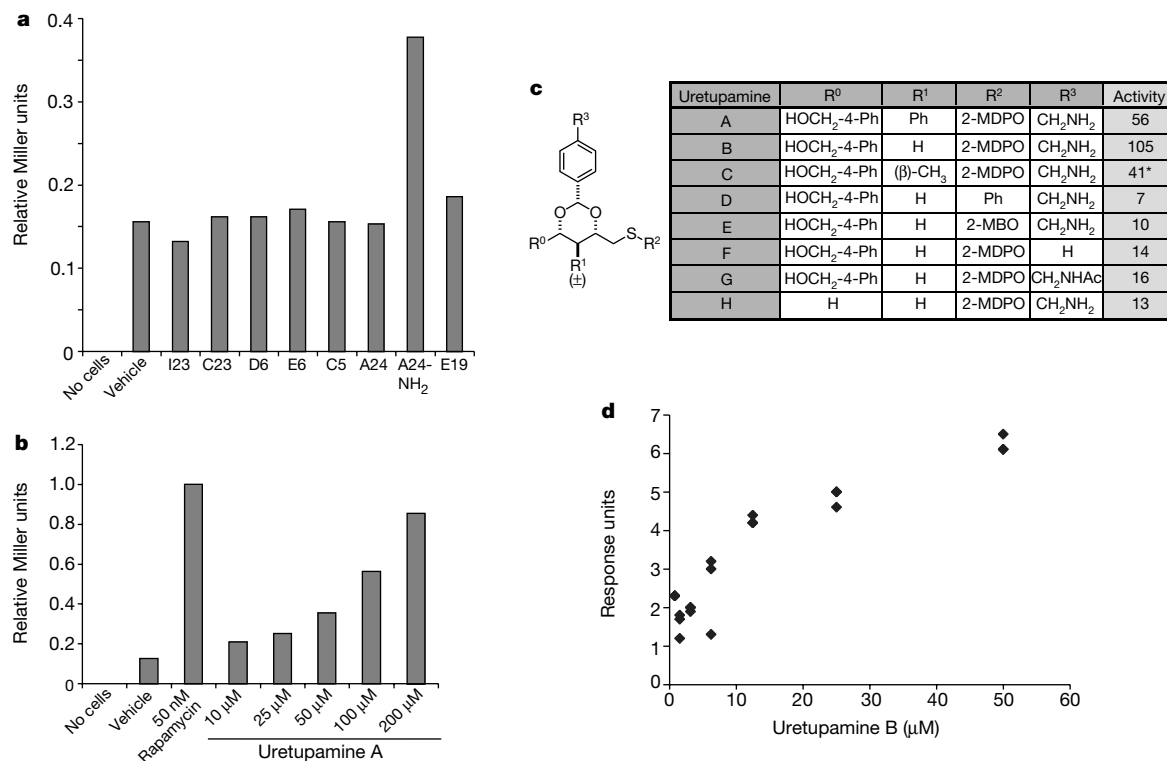


Figure 2 Studies *in vivo*, dose-response and structure-activity relationships of uretupamine. **a**, A yeast strain (DB26-3A) growing in YPD medium expressing a *PUT1-lacZ* reporter was treated with 50 nM rapamycin or with a compound that had been detected to bind to labelled Ure2p on a small-molecule microarray. After 90 min of treatment at 30 °C, a standard liquid β -galactosidase assay was performed. Data are expressed in fold Miller units compared with treatment with 50 nM rapamycin for 90 min. DB26-3A (*MAT α ura3-52 ade2 PUT1-lacZ*) was a gift from M. Brandriss. Vehicle: samples treated with *N,N*-dimethylformamide (DMF), the vehicle into which library compounds were dissolved. **b**, Uretupamine A was resynthesized and tested in the β -galactosidase assay by using the *PUT1-lacZ* reporter at the concentrations indicated for 60 min at 30 °C in YPD medium. **c**, Compounds derived from the uretupamine A structure were synthesized to explore structure-activity relationships. Listed are β -galactosidase

assay results of treatment with each compound at 100 μ M (asterisk designates 50 μ M) for 60 min at 30 °C in YPD medium. Data are in percentage Miller units compared with treating with 50 nM rapamycin for 60 min. Ac, acetate; MDPO, 2-mercapto-4,5-diphenyloxazole; MBO, 2-mercaptobenzoxazole; Ph, phenyl. **d**, Binding of uretupamine B to Ure2p was determined by using surface plasmon resonance (BIAcore 3000) to have a dissociation constant of 7.5 μ M. Data points were acquired in triplicate. Ure2p was immobilized to CM5 sensor chips by injection of 100 $\mu\text{g ml}^{-1}$ Ure2p in 10 mM sodium acetate pH 4.5 in accordance with the manufacturer's procedures. The reference cell was derivatized with antibodies against glutathione *S*-transferase (GST) followed by GST capture. Small-molecule binding measurements and dissociation were in PBS/Tween-20 containing 10% DMF flow rate 5 $\mu\text{l min}^{-1}$).

(Fig. 2c). As expected, uretupamine B was more potent than uretupamine A (Fig. 2c). Surface plasmon resonance was used to obtain a binding constant for uretupamine A and B binding to purified Ure2p. This demonstrated that uretupamine A and B bound to Ure2p with equilibrium dissociation constants of 18.1 and 7.5 μ M, respectively (Fig. 2d), which is consistent with their potencies in cells.

To determine the precise effects and specificity of uretupamine, we used whole-genome transcription profiling in wild-type cells as well as an otherwise isogenic *ure2 Δ* strain—a ‘targetless’ strain¹⁸. (Complete transcription profiling data are publicly available in Supplementary Information and at <http://www.schreiber.chem.harvard.edu>.) Both uretupamine A and B upregulated only a subset of genes (including *PUT1*, *PUT2*, *PRB1*, *NIL1* and *UGA1*) known to be under the control of Ure2p (Fig. 3a, b). The expression of other genes (including *GAP1*, *MEP2*, *AGP1*, *BAT2* and *DAL5*) controlled by Ure2p was essentially unchanged (Fig. 3a, b). The compounds had little or no effect on either set of genes in a targetless *ure2 Δ* strain, an otherwise identical strain lacking only the gene encoding the putative protein to which uretupamines A and B bind (Fig. 3a, b). This result suggests that a small molecule readily obtained from diversity-oriented synthesis and screening with the use of small-molecule microarrays has nearly complete cellular specificity for its screening partner, at least as judged by its global effects on the mRNA levels of treated cells.

Although Ure2p-controlled genes are normally thought of as responsive to nitrogen quality, the subset of genes induced by uretupamine (*PUT1*, *PUT2*, *PRB1*, *NIL1* and *UGA1*) has been shown to be upregulated when glucose is removed from the media¹⁹. The mechanism for this differential regulation of Ure2p-controlled genes in response to different nutrient signals is not understood. Ure2p represses transcription factors Gln3p and Nil1p, which might be differentially regulated to achieve this effect^{19,20}. To test this hypothesis, we therefore profiled uretupamine B in *gln3 Δ* and *nil1 Δ* strains. Remarkably, we found that deleting *GLN3* had

little effect on the actions of uretupamine, whereas deleting *NIL1* abrogated its actions (Fig. 3b). Further confirmation for this selectivity comes from whole-genome vector-based comparisons^{19,21} of four profiles; these comparisons show that *URE2* and *NIL1*, but not *GLN3*, are critical for uretupamine action (Fig. 3c). Northern blot analysis confirmed the effect of uretupamine B on *PUT1* expression (normalized to *ACT1* expression) in wild-type, *gln3 Δ* and *nil1 Δ* strains (data not shown).

The fact that the binding of uretupamine to Ure2p induces the expression of glucose-sensitive genes in a *NIL1*-dependent manner suggests that Ure2p might itself be the target of a glucose-sensitive pathway. This is in contrast to a glucose-sensitive pathway impinging on Nil1p, bypassing Ure2p.

Because Ure2p is a phosphoprotein^{9,10,22}, we examined the phosphorylation state of Ure2p after different types of nutrient shift. We shifted wild-type cells from the high-quality nitrogen source, ammonium sulphate, to the low-quality nitrogen source proline. We also shifted cells from the high-quality (fermentable) energy source glucose to the low-quality (non-fermentable) energy sources acetate or glycerol. Surprisingly, Ure2p was not dephosphorylated when ammonium sulphate was removed but was dephosphorylated when glucose was removed (Fig. 4a). These data indicate that signals not previously thought to regulate Ure2p alter its phosphorylation state, whereas signals previously thought to regulate Ure2p do not alter its phosphorylation state. It remains to be seen whether there is another means by which these signals can regulate Ure2p function. Identical results were obtained for cells transferred from a glucose-containing medium to an ethanol-containing medium and for cells of a different background (W303) (data not shown). Other stresses (such as 1 M NaCl, 1 M sorbitol, pH 9.5 or heat shock) known to upregulate Ure2p-dependent genes^{23,24} did not cause Ure2p dephosphorylation (Fig. 4b, data not shown). These data suggest that Ure2p is part of a signalling pathway that specifically responds to glucose.

Kornberg and Krebs first proposed that on energy sources such as

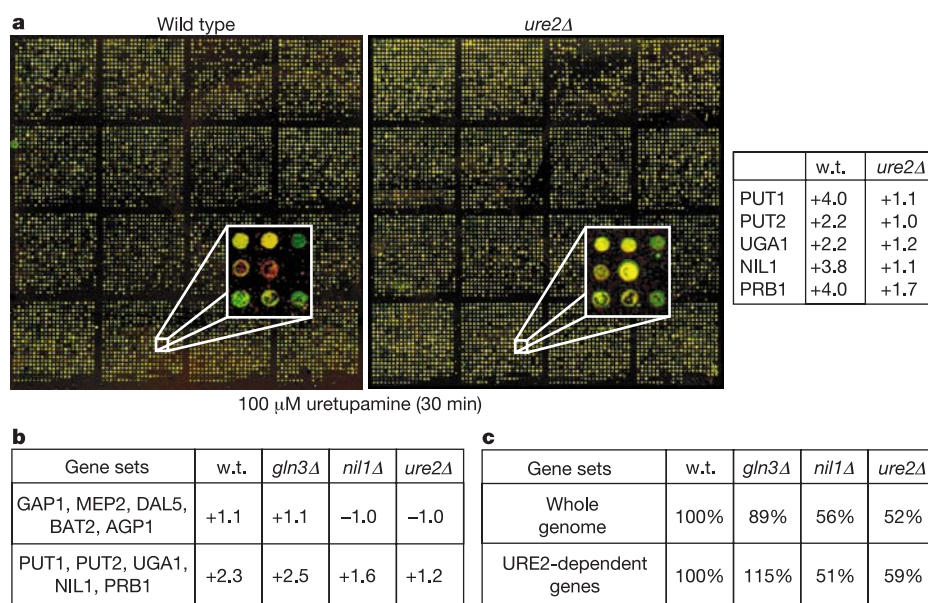


Figure 3 Transcription profiling of treatment with uretupamine. **a**, The left microarray corresponds to wild-type cells (PM38) treated with vehicle (DMF) versus wild-type (w.t.) cells treated for 30 min with uretupamine A at 100 μ M. The right microarray corresponds to *ure2 Δ* cells (PH2) treated with vehicle versus *ure2 Δ* cells treated for 30 min with uretupamine A at 100 μ M. Profiles were obtained as described⁹. At the right are shown specific gene inductions of some *URE2*-dependent genes from the microarrays. PM38 (*MAT α leu2-3,112 ura3-52*), PM71 (*MAT α leu2-3,112 ura3-52 gln3 Δ 5::LEU2*), MS221

(*MAT α ura3-52 nil1::hisG*), PH2 (*MAT α leu2-3,112 ura3-52 ure2 Δ 12::URA3*) were gifts from B. Magasanik and M. Brandriss. **b**, The transcription profiles of wild-type cells (PM38), *ure2 Δ* cells (PH2), *gln3 Δ* cells (PM71) and *nil1 Δ* cells (MS221) grown in YPD medium treated for 30 min with 100 μ M uretupamine B were obtained. The geometric means of gene inductions of the sets listed are shown. **c**, An analysis was performed based on treating individual profiles as high-dimensional vectors and then examining the ratios of their magnitudes as a measure of relative activity^{19,21}.

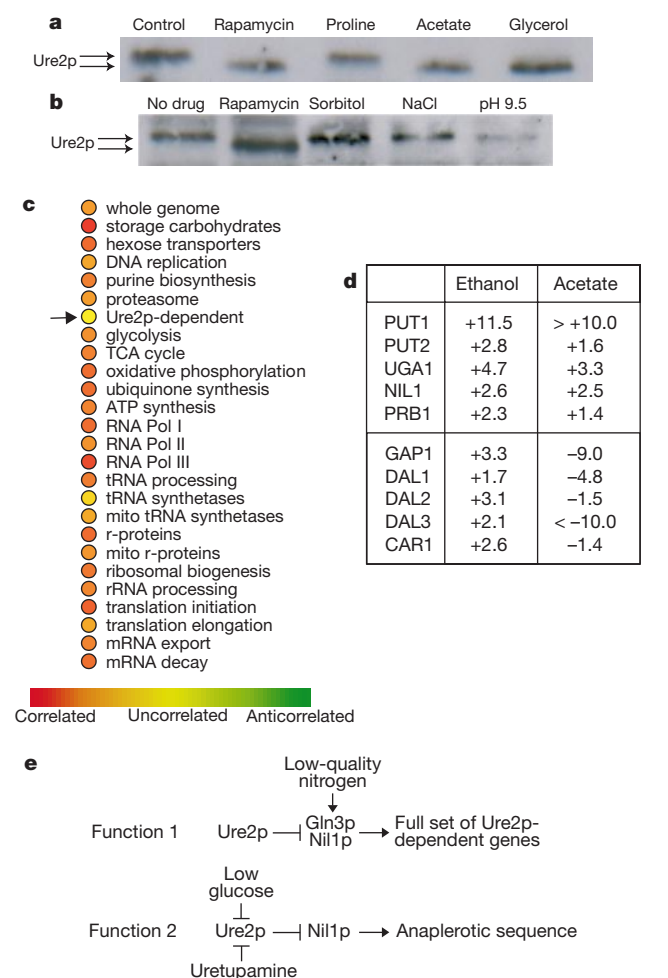


Figure 4 Glucose-sensitive signalling and a model of Ure2p function. **a**, A wild-type strain (PM38) was grown to mid-exponential phase in synthetic glucose (dextrose) (SD)-ammonium sulphate (AS) medium, washed with PBS and split into either SD-AS medium, SD-AS medium containing 50 nM rapamycin, SD-proline medium, synthetic acetate-AS medium or synthetic glycerol-AS medium, then incubated with shaking for 1 h at 30 °C. Control: the sample split into SD-AS medium. Synthetic medium consisted of 1.7 g of YNB medium, without amino acids and without AS, 2% carbon source, 0.1% nitrogen source, and auxotrophic supplements when needed (leucine 120 mg l⁻¹, uracil 20 mg l⁻¹). Whole-cell lysates were blotted with anti-Ure2p antibodies kindly provided by R. Wickner as described previously¹⁹. A previous report claimed that Ure2p-dephosphorylation does occur upon nitrogen limitation but those authors examined cells shifted from a rich medium to a synthetic-nitrogen-limited medium, thus changing many variables of the medium simultaneously¹⁰. In our experiments, cells were shifted from a synthetic medium containing 0.1% ammonium sulphate as a high-quality nitrogen source to an otherwise identical medium containing 0.1% proline as a low-quality nitrogen source. Inspection of published Ure2p immunoblots under similar conditions also supports the lack of a mobility shift^{27,28}. **b**, A wild-type strain (PM38) was treated for 1 h with no drug, 50 nM rapamycin, salt stress (1 M NaCl), high osmolarity (1 M sorbitol) or high pH (pH 9.5). **c**, A wild-type strain (PM38) was shifted from an SD-AS medium to a synthetic acetate-AS medium and incubated with shaking for 30 min at 30 °C. This transcription profile was compared to the profile of the same strain shifted from SD-AS to synthetic ethanol-AS medium¹⁹. The expression of various subsets of genes was compared between profiles by using vector algebra and the results are presented in a colorimetric comparison array^{19,21}. **d**, Inspection of Ure2p-dependent genes reveals that uretupamine-sensitive genes behave similarly when cells are shifted to ethanol or acetate, whereas uretupamine-insensitive genes behave differently. **e**, Nitrogen quality regulates the Ure2p-Gln3p/Nil1p complex by signalling to Gln3p/Nil1p, whereas low glucose concentration regulates the complex by signalling to Ure2p. Low glucose concentration leads to dephosphorylation of Ure2p to induce a specific set of genes involved in an anaplerotic sequence, a state mimicked by uretupamine.

acetate, metabolic sequences called anaplerotic are activated to replenish tricarboxylic-acid-cycle intermediates²⁵. Yeast cells growing in acetate-containing media have been shown to accumulate ammonia²⁶, which leads to the following paradox. Ammonia would repress the expression of Ure2p-dependent genes, including those thought to promote survival on acetate as part of an anaplerotic sequence (*PUT1*, *PUT2* and *UGA1*)²⁰. It is possible that acetate-induced Ure2p-dephosphorylation protects the anaplerotic sequence from this repression by ammonia. We performed the transcription profile of yeast shifted from glucose to acetate and compared it with that of cells shifted from glucose to ethanol. Genome-wide analysis showed that Ure2p-dependent genes were differentially affected by the two transitions (Fig. 4c). Unlike ethanol, acetate caused the downregulation of some Ure2p-dependent genes but, like ethanol, acetate induced those genes activated by uretupamine (Fig. 4d). Taken together, these data suggest that Ure2p-dephosphorylation stabilizes the induction of genes for an anaplerotic sequence when other cellular forces might repress the expression of these same genes (Fig. 4e).

Our approach to uncovering the role of Ure2p in glucose signalling is rooted in the principles of reverse genetics. We desired a method of modulating Ure2p function selectively to examine the resulting phenotype. Because uretupamine modulates only a subset of Ure2p function, its effects are more specific than those resulting from deletion of the *URE2* gene. This property of uretupamine highlights the multifunctionality of individual proteins and addresses the challenge in proteomics to identify and control all possible inputs and outputs of each protein. With uretupamine, we demonstrated a functional connection between Ure2p, Nil1p and glucose levels. We additionally have a means to control this system more selectively than any physiological stimulus or genetic deletion. Diversity-oriented synthesis and small-molecule microarrays provide a potentially systematic method for acquiring powerful probes, where different small molecules can modulate different aspects of a protein's function, preceding the discovery of a genetic allele of a similar phenotype. □

Received 13 December 2001; accepted 5 February 2002.

- Blinder, D., Coschigano, P. W. & Magasanik, B. Interaction of the GATA factor Gln3p with the nitrogen regulator Ure2p in *Saccharomyces cerevisiae*. *J. Bacteriol.* **178**, 4734–4736 (1996).
- Beck, T. & Hall, M. N. The TOR signalling pathway controls nuclear localization of nutrient-regulated transcription factors. *Nature* **402**, 689–692 (1999).
- Cunningham, T. S., Andhare, R. & Cooper, T. G. Nitrogen catabolite repression of DAL80 expression depends on the relative levels of Gat1p and Ure2p production in *Saccharomyces cerevisiae*. *J. Biol. Chem.* **275**, 14408–14414 (2000).
- Wiemann, S. *et al.* Toward a catalog of human genes and proteins: sequencing and analysis of 500 novel complete protein coding human cDNAs. *Genome Res.* **11**, 422–435 (2001).
- Narahashi, T., Moore, J. W. & Scott, W. R. Tetrodotoxin blockage of sodium conductance increase in lobster giant axons. *J. Gen. Physiol.* **47**, 965–974 (1964).
- Lehmann, J. M. *et al.* An antidiabetic thiazolidinedione is a high affinity ligand for peroxisome proliferator-activated receptor γ (PPAR γ). *J. Biol. Chem.* **270**, 12953–12956 (1995).
- Coschigano, P. W. & Magasanik, B. The URE2 gene product of *Saccharomyces cerevisiae* plays an important role in the cellular response to the nitrogen source and has homology to glutathione S-transferases. *Mol. Cell. Biol.* **11**, 822–832 (1991).
- Wickner, R. B. [URE3] as an altered URE2 protein: evidence for a prion analog in *Saccharomyces cerevisiae*. *Science* **264**, 566–569 (1994).
- Hardwick, J. S., Kuruvilla, F. G., Tong, J. K., Shamji, A. F. & Schreiber, S. L. Rapamycin-modulated transcription defines the subset of nutrient-sensitive signaling pathways directly controlled by the Tor proteins. *Proc. Natl Acad. Sci. USA* **96**, 14866–14870 (1999).
- Cardenas, M. E., Cutler, N. S., Lorenz, M. C., Di Como, C. J. & Heitman, J. The TOR signaling cascade regulates gene expression in response to nutrients. *Genes Dev.* **13**, 3271–3279 (1999).
- Sternson, S. M., Louca, J. B., Wong, J. C. & Schreiber, S. L. Split-pool synthesis of 1,3-dioxanes leading to arrayed stock solutions of single compounds sufficient for multiple phenotypic and protein-binding assays. *J. Am. Chem. Soc.* **123**, 1740–1747 (2001).
- Schreiber, S. L. Target-oriented and diversity-oriented organic synthesis in drug discovery. *Science* **287**, 1964–1969 (2000).
- Blackwell, H. E. *et al.* A one-bead, one-stock solution approach to chemical genetics: part 1. *Chem. Biol.* **8**, 1167–1182 (2001).
- Clemons, P. A. *et al.* A one-bead, one-stock solution approach to chemical genetics: part 2. *Chem. Biol.* **8**, 1183–1195 (2001).
- MacBeath, G., Koehler, A. N. & Schreiber, S. L. Printing small molecules as microarrays and detecting protein–ligand interactions en masse. *J. Am. Chem. Soc.* **121**, 7967–7968 (1999).
- Hergenrother, P. J., Depew, K. M. & Schreiber, S. L. Small-molecule microarrays: covalent attachment and screening of alcohol-containing small molecules on glass slides. *J. Am. Chem. Soc.* **122**, 7849–7850 (2000).

17. Xu, S., Falvey, D. A. & Brandriss, M. C. Roles of URE2 and GLN3 in the proline utilization pathway in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **15**, 2321–2330 (1995).
18. Marton, M. J. *et al.* Drug target validation and identification of secondary drug target effects using DNA microarrays. *Nature Med.* **4**, 1293–1301 (1998).
19. Shamji, A. F., Kuruvilla, F. G. & Schreiber, S. L. Partitioning the transcriptional program induced by rapamycin among the effectors of the Tor proteins. *Curr. Biol.* **10**, 1574–1581 (2000).
20. Kuruvilla, F. G., Shamji, A. F. & Schreiber, S. L. Carbon- and nitrogen-quality signaling to translation are mediated by distinct GATA-type transcription factors. *Proc. Natl Acad. Sci. USA* **98**, 7283–7288 (2001).
21. Kuruvilla, F. G., Park P. J. & Schreiber, S. L. Vector algebra in the analysis of genome-wide expression data. *Genome Biol.* **3**(3), 0011.1–0011.11 (2002).
22. Bertram, P. G. *et al.* Tripartite regulation of Gln3p by TOR, Ure2p and phosphatases. *J. Biol. Chem.* **275**, 35727–35733 (2000).
23. Causton, H. C. *et al.* Remodeling of yeast genome expression in response to environmental changes. *Mol. Biol. Cell* **12**, 323–337 (2001).
24. Gasch, A. P. *et al.* Genomic expression programs in the response of yeast cells to environmental changes. *Mol. Biol. Cell* **11**, 4241–4257 (2000).
25. Kornberg, H. L. & Krebs, H. A. Synthesis of cell constituents from C₂ units by a modified tricarboxylic acid cycle. *Nature* **179**, 988–991 (1957).
26. Bogonez, E., Machado, A. & Satrustegui, J. Ammonia accumulation in acetate-growing yeast. *Biochim. Biophys. Acta* **733**, 234–241 (1983).
27. Edskes, H. K., Hanover, J. A. & Wickner, R. B. Mks1p is a regulator of nitrogen catabolism upstream of Ure2p in *Saccharomyces cerevisiae*. *Genetics* **153**, 585–594 (1999).
28. Edskes, H. K. & Wickner, R. B. A protein required for prion generation: [URE3] induction requires the ras-regulated mks1 protein. *Proc. Natl Acad. Sci. USA* **97**, 6625–6629 (2000).

Supplementary Information accompanies the paper on Nature's website (<http://www.nature.com>).

Acknowledgements

We thank R. Melki for providing bacterially expressed Ure2p protein. F.G.K. was supported by the NIH Medical Scientist Training Program, A.E.S. by the Howard Hughes Medical Institute predoctoral fellowship, S.M.S. by the Roche and NSF predoctoral fellowships, and P.J.H. by the American Cancer Society. S.L.S. is an Investigator at the Howard Hughes Medical Institute. This research was funded by a grant from the NIGMS (GM-38627).

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to S.L.S. (e-mail: sls@slsiris.harvard.edu).

A 'periodic table' for protein structures

William R. Taylor

Division of Mathematical Biology, National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK

Current structural genomics programs aim systematically to determine the structures of all proteins coded in both human and other genomes, providing a complete picture of the number and variety of protein structures that exist. In the past, estimates have been made on the basis of the incomplete sample of structures currently known. These estimates have varied greatly (between 1,000 and 10,000; see for example refs 1 and 2), partly because of limited sample size but also owing to the difficulties of distinguishing one structure from another. This distinction is usually topological, based on the fold of the protein; however, in strict topological terms (neglecting to consider intra-chain cross-links), protein chains are open strings and hence are all identical. To avoid this trivial result, topologies are determined by considering secondary links in the form of intra-chain hydrogen bonds (secondary structure) and tertiary links formed by the packing of secondary structures. However, small additions to or loss of structure can make large changes to these perceived

topologies and such subjective solutions are neither robust nor amenable to automation. Here I formalize both secondary and tertiary links to allow the rigorous and automatic definition of protein topology.

The organization and classification of the bewildering variety of protein structure has been approached using clustering methods. Various computer programs have been devised to measure the three-dimensional similarity of one protein coordinate set to another³. From these measures, similar proteins can be grouped together, given a name, and arranged in a hierarchical clustering with others that share some partial or overall similarity. Depending on the method of comparison, or the extent of expert judgements needed, differing classifications of protein structure have emerged, ranging from one that is almost completely expert-based⁴, through partially automated methods^{5,6} to an almost fully automatic method⁷. The drawback of these hierarchical approaches is that, although the close relationships between similar proteins are reasonably well defined, the more tentative relationships that give the large-scale structure to the hierarchy are usually beyond the ability of the computer programs to recognize and are subject to variation when defined by experts. As a result, the classifications tend to have numerous small clusters (families of super families) all roughly grouped into just a few categories that are based on overall secondary structure content and arrangement.

An important secondary problem in the classification of proteins is how large proteins should be divided into pieces (domains) that can be classified more easily. The current approach among the more automated methods is first to divide and then to classify. However, it is clear from the expert-based approach that the initial process of classification can affect how the protein is then broken into domains and so differences in domain definition are a major source of inconsistency among the current classification systems⁸.

To avoid the problems associated with a hierarchy, the method outlined here is based on a set of idealized structures that are compared with all known structures. (The programs and data described in this work can be found at: <http://mathbio.nimr.mrc.ac.uk/ftp/wtaylor>.) The domain definition problem is less directly solved, although as the ideal structures are all of domain size, the best match can be used to define (or bias) the definition of the domains. This approach is unusual in that it shifts the classification from a clustering problem to that of finding the best set of ideal structures that can account for as much protein structure as possible. As the ideal structures will be generated from rules applied to basic 'Forms', this can be viewed as finding a minimum basis set of generating Forms. These Forms were derived from a model in

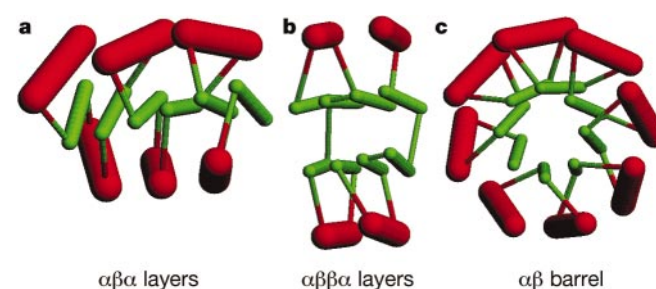


Figure 1 Stick-figure representations of the basic Forms. Each of the basic generating Forms is represented by 'stick' models in which α -helices are red and drawn thicker than the green β -strands. **a**, $\alpha\beta$ layers. Six strands are shown, but the sheet can extend indefinitely. **b**, $\alpha\beta\alpha$ layers. As in **a**, the sheets can be extended. (Removal of the α -layers leaves the common β -'sandwich'). **c**, Eight-fold $\alpha\beta$ barrel. Similar barrels with 5–9 strands were constructed. (See Supplementary Information A.1 for construction details). By deleting helices and strands from these models, almost all known globular protein domains of β and $\beta\alpha$ types can be generated. Figures 1 and 4 were prepared using the program RasMol (<http://www.umass.edu/microbio/rasmol>).

which the hydrogen-bonded links across a β -sheet impose a layer structure onto the arrangement of secondary structures in a protein domain (Fig. 1)^{9,10}. These layers can consist of either α -structure (packed α -helices) or β -structure (hydrogen-bonded β -strands). There are seldom more than four layers in any one domain and each layer tends to be exclusively composed of one of these two types of secondary structure. The spacing between the axes of packed α -helices is typically 10 Å, as is the spacing between β -sheets and between helices and sheets^{11,12}, whereas the spacing between the hydrogen-bonded strands in a sheet is close to 5 Å. This makes 10 Å a convenient unit with which to 'digitize' protein structure.

β -sheets normally have a twist, and the whole structure follows this twist, resulting in a staggered arrangement for the secondary structure elements in the outer layers. Sheets can also curl and incorporate a stagger between adjacent strands—combinations of these 'distortions' can result in the sheet forming a complete hydrogen-bonded cylinder (referred to as a barrel)¹³. All these parameters (twist, curl, stagger) might be varied with different numbers of layers (of different composition), but as a simple beginning, the more limited layer combinations shown in Fig. 2 were considered and the curl and stagger co-varied to give either topologically flat sheets or cylindrical sheets (barrels), with the curled sheets being represented by partial barrels. The resulting arrangement is not dissimilar to the periodic table of elements. In this loose analogy, the layers are equivalent to valence shells that become progressively filled with electrons (secondary structure elements): first, the inner β -layer (s orbitals) followed by the outer α layers (p orbitals) then repeating with a second β -layer (d orbitals). Extending the model to incorporate the permutations arising from additional α -layers would be reminiscent of the intersection of the rare-earth series into the periodic table. (See

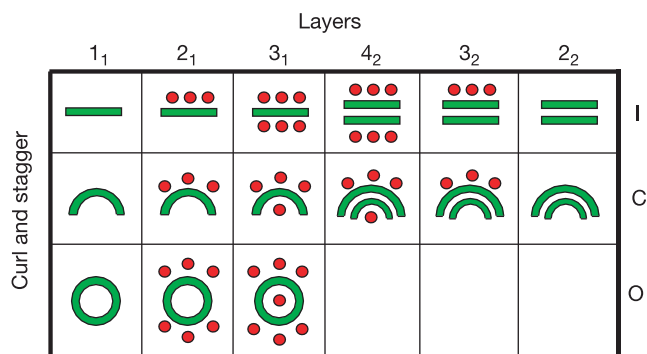


Figure 2 Simplified layer structure of proteins. Layers of secondary structure (β , green; α , red) are combined to make globular protein domains. The β -sheets are represented as bars and circles, as they would appear when viewed looking along their component strands. Each sheet has a left-handed twist between the strands (not depicted) onto which can be added curl and stagger. This allows the sheets to progressively 'deform' from a topologically flat sheet into a cylinder (or barrel). The two endpoints and one intermediate stage are represented by the rows in the figure and indexed as I (flat), C (curled) and O (barrel). For each of these, up to four layers of secondary structure are shown (α -helices are drawn as red dots viewed end-on). For simplicity, not all possible layer combinations have been represented; in particular, those with adjacent α -layers have been omitted because the boundary between these is not well defined. Most biologically important structures can be generated from three of those represented above: this set of three is referred to as the basis set. (See Fig. 1 for three-dimensional 'stick' figures of the basis set.) Using the I, C, O index plus layer number: I3₁ can generate I2₁ and I1₁ (by the deletion of helices); O2₁ can generate O1₁, and with the removal of strands from the barrel, also C2₁ and C1₁. Similarly, I4₂ can produce an $\alpha\beta\beta$ and the common $\beta\beta$ layer structures. Of those remaining, only C3₁ is biologically important and will be reconsidered later. All possible deletions are made for each basic structure: those for O2 are shown in Fig. 3.

Supplementary Information A.3 for further discussion). Because there is no strict energetic difference between different protein structures, the filling of the layers with secondary structures was allowed more freedom than their electronic counterparts and, so for each of the basic Forms described above, any combination of secondary structures can be removed. For example; an eight-stranded barrel with eight helices around it can give rise to the variations shown in Fig. 3.

In order to match all ideal structures against all known protein structures, each native structure was first reduced to linear segments¹⁴. For each comparison, a fast bipartite graph-matching algorithm²¹ was used as a pre-filter and also to 'prime' a more exhaustive double-dynamic programming comparison algorithm¹⁵. Each comparison began by pairing up line segments ('sticks') irrespective of their length or direction, but when a good fit was found, the lengths of the native protein 'sticks' were set to 10 Å and the connectivity of the ideal sticks were set to match the native protein. This allowed the two matching stick-structures to be directly superposed in three dimensions using a conventional comparison program and the quality of the match was taken as the root-mean-squared deviation (r.m.s.d.) between two sets of stick endpoints. From examination of the solutions, anything less than 5 Å r.m.s.d. between two structures is a good fit, and between 5 and 6 Å r.m.s.d. is acceptable. An example of two solutions in these ranges is shown in the Supplementary Information A.1.

The method was initially tested on the single domains defined to have distinct folds in the CATH database⁵. From these, the all- α

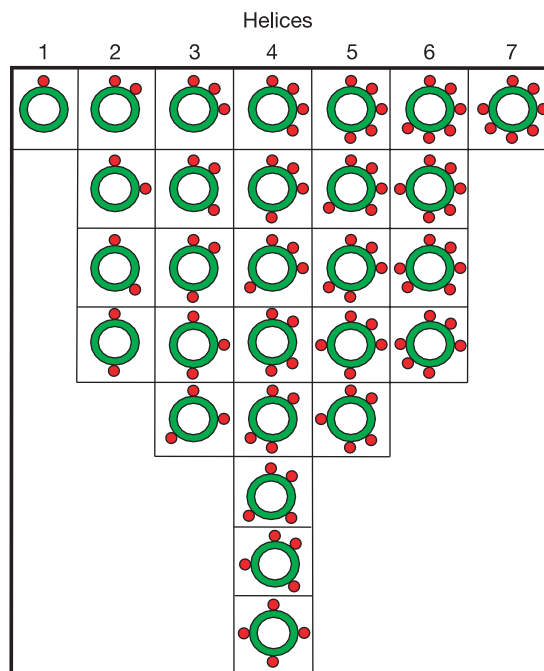


Figure 3 β -barrel helix packing variations. The basic barrel Form (O2 in Fig. 2), with eight strands and eight helices, can generate 30 packing variations through the deletion of helices (28 depicted above plus the two not shown with eight and zero helices). A similar combinatorial enumeration of variants was made for barrels of 5–9 strands and flat sheets of 3–13 strands. The intermediate curled strands (row C in Fig. 2) were made by successive deletion of all but three strands from each of the barrels (and the helices similarly permuted). The application of symmetry considerations greatly reduced the number of possible combinations, but for some larger structures it was necessary to impose a limit. Variations on the large 'flat' layers were ranked by compactness and (for alphabetic reasons) just the 26 most compact combinations were considered for matching. Variations on the larger barrels were also limited by allowing only one break in the sheet layer and similarly restricted to 26 variants on any given combination of secondary structures. In total, 12,640 variations were generated.

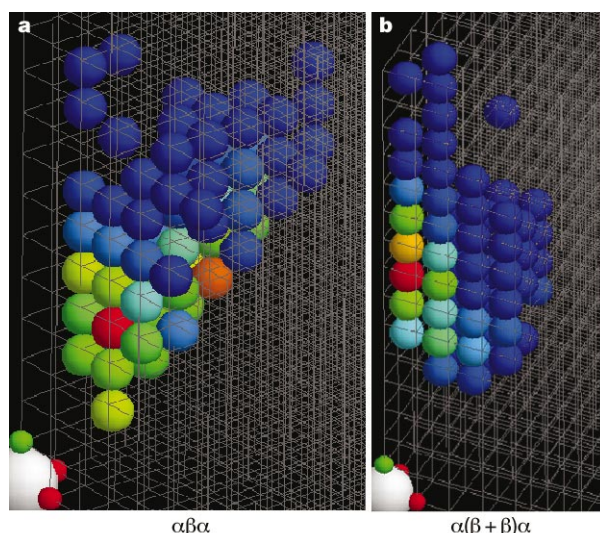


Figure 4 Structure tables for flat sheets. The number of helices and strands in each secondary structure layer of the ideal Forms is represented as a grid: **a**, $\alpha\beta\alpha$; and **b**, $\alpha(\beta+\beta)\alpha$. The large white sphere marks the origin of the grid with the directions of increasing β and α indicated by small green and red attached spheres. (For the $\alpha\beta\beta\alpha$ structures, the number of β -strands in each sheet have been combined). Grid cells contain a sphere if their corresponding ideal structure constitutes the best match to a native protein domain. This is coloured to indicate the number of proteins that match (red, most; blue, least). In **a**, the most populated cell has the Form 0-4-2 (α - β - α) with 48 members, closely followed by the 2-5-3 Form (42 members). The former included 25 different topologies while the latter includes only 18 (being dominated by the common flavodoxin fold). The most populated cell in the $\alpha\beta\beta\alpha$ table (83 members) has seven strands and no helices, corresponding to the 3-on-4 β -sandwich of the immunoglobulin fold. The table for the barrel structures (not shown) has a single highly populated cell for the $(\beta\alpha)_8$ barrel (8-8-8).

domains were excluded, as were integral membrane proteins. Although ideal structures can be described for these classes^{16,17}, a comprehensive nomenclature has not been devised. Some small domains that do not have packed secondary structure were also excluded, leaving a sample of 418 domains. Of these domains, only 34 had no acceptable fit to an instance of the ideal Forms, and most of these were small, with fragmentary beta structure. Of the remaining 384 domains 85% of the matches accounted for over 50% of the structure with more than 70% of the structure being matched in over half of the domains. Partially matched domains comprised large β -propeller structures and viral coat proteins that have long 'unstructured' loops. The remainder was typically composed of α -helices which were either packed in a distinct subdomain or suggested an additional α -layer. Rather than adapt the current model to account for these elements, it was clear that it would first be necessary to reassess the definition of domains. This was done using an automatic (Ising-like) method¹⁸ in which it was possible to bias the matched portion of a protein to remain distinct from the rest. Any remaining material was then presented again for matching to the ideal structures and the process repeated. The full chains from a non-redundant set of 2,230 protein structures were then processed by this algorithm, resulting in an improved coverage by the ideal Forms which, on average, now accounted for 80% of each protein, with 70% of the structure being matched in 75% of the domains (previously just over 50%).

Each Form can be indexed by the number of secondary structures in each layer (see Supplementary Information A.2 for details) and from the index of its matching Form, each domain can be allocated a grid reference and plotted in space. Any step in the planar- β grids represents the addition or removal of a secondary structure, and one of the dimensions in the barrel grid represents the opening of the barrel (for a fixed number of strands). Using this representation, the full range of known protein structure can be visualized (Fig. 4).

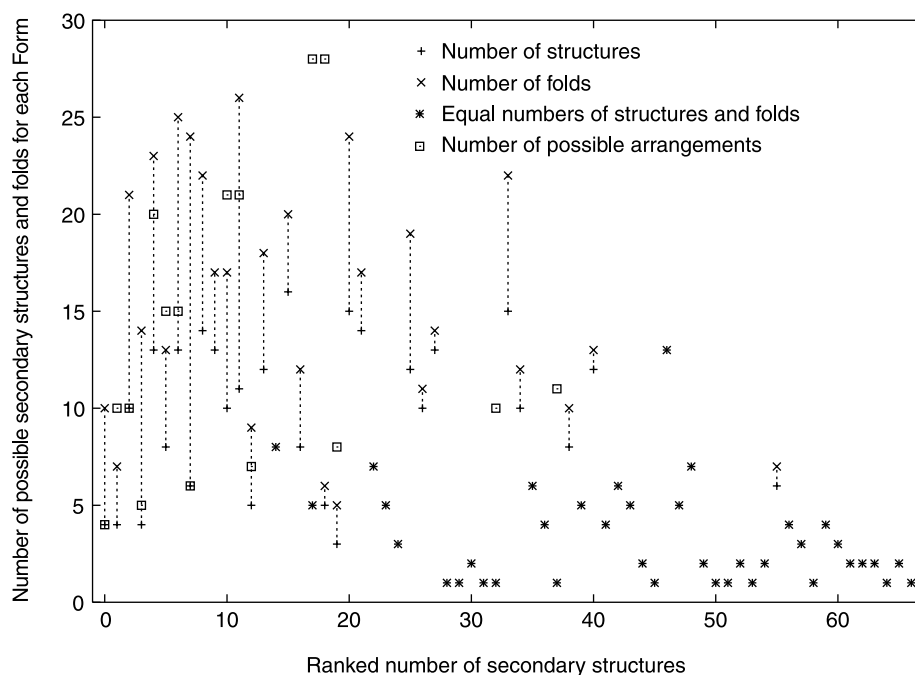


Figure 5 Fold numbers for ranked Forms. The three-layer Forms (Fig. 4a) were ranked by the number of secondary structures they contain. Those of equal size were ordered by the number of β , then α elements. Against this is plotted the number of different secondary structure arrangements seen for each Form (upright crosses) and the number of different folds (tilted crosses). (See Supplementary Information A.2 for details). The number of

observed arrangements and the number of folds are linked by a dashed line and when they are equal, appear as an asterisk. For reference, the number of possible arrangements is plotted as a box but this only appears for some of the smaller Forms as, for most, it is a very large number.

This classification of structure has not explicitly considered topology but has concentrated on the prerequisite of defining the secondary and tertiary links. From this basis, the previously difficult issue of topology becomes almost trivial, because two proteins matching the same ideal Form will either have the same connection of their secondary structures or will not. (See Supplementary Information A.2 for nomenclature details). Topology descriptions were written for each protein in each grid-cell and sorted for uniqueness, thus giving the number of folds in each cell. This value was then plotted along with the number of unique secondary structure strings for the three-layer Forms (Fig. 5). This simple analysis shows that for the smallest Forms, all linear arrangements of secondary structure have been observed and there are typically twice as many folds as secondary structure arrangements. With over ten secondary structures (around 35 in the ranked Forms in Fig. 5), this balance changes, and almost every secondary structure arrangement corresponds to a unique fold. This suggests that with these larger $\beta\alpha$ proteins, sequence-structure matching (threading) methods¹⁹ need only concentrate on the correct prediction of secondary structure to find an unknown fold and not on their three-dimensional interaction.

If a structure matches, say, a 2-5-3 (α - β - α) Form then it will also match any substructure at least as well (2-5-2, 1-5-3, 2-4-2...). and two proteins which have a different topology when matched against the 2-5-3 Form might have the same core 2-5-2 topology. The question of whether two proteins have the same fold now becomes relative and should be posed thus: what is the largest ideal Form under which two proteins have the same fold? Although the largest Form with common topology will be of greatest interest, it is also informative to view all the subsidiary matches in the format of the grids used in Fig. 4. Such a representation allows not only pairs, but also whole groups of proteins to be analysed and it is simple to determine both visually and automatically if they share a common core. Each step through the grid from one protein to another (via subsidiary matches) represents a deletion or addition of a secondary structure element. This allows a path to be defined between any two structures in the same grid (no matter how dissimilar) and a minimum path length to be computed.

Besides the all- α class of protein, the only structures not 'captured' by the set of three basic Forms used above were β proteins that contain internal repetition. These included not only the series of propeller proteins but also those with a triangular arrangement of structure, including β -trefoils, β -prisms and β -helices²⁰. Clearly triangles do not map well onto layers (or cylinders) and the only solution for these may be to generate a 'triangular' version of Fig. 2, or else to treat them as exceptions because of their detectable internal sequence repeats. For the moment, the latter route will be followed but a series of propeller Forms will be added to the basis set, along with the neglected C3₁ Form from Fig. 2 (which includes the β -grasp structures).

Thus, by breaking down structures into their basic Forms I have opened up a new approach to the classification and analysis of protein structure that is both flexible and automatic. Unlike clustering methods, it does not require the comparison of all structures to each other and can therefore be incrementally updated as new structures are determined. The major result from this approach is that it allows the difficult problem of protein topology to be rigorously addressed. The fold of any individual protein can be specified as that of the largest matching Form, and for pairs or groups of proteins, the fold is that of the largest common Form. This implies that the topology of a protein can only be defined under the specification of a given ideal Form which, in turn, means that proteins do not have a unique intrinsic topology or a unique position in any classification that is based on topology. □

Received 12 October 2001; accepted 28 January 2002.

1. Chothia, C. One thousand families for the molecular biologist. *Nature* **357**, 543–544 (1992).

- Orengo, C. A., Jones, D. T. & Thornton, J. Protein superfamilies and domain superfolds. *Nature* **372**, 631–634 (1994).
- Eidhammer, L., Jonassen, I. & Taylor, W. R. Structure comparison and structure patterns. *J. Comput. Biol.* **7**, 658–716 (2000).
- Murzin, A. G., Brenner, S. E., Hubbard, T. & Chothia, C. SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J. Mol. Biol.* **247**, 536–540 (1995).
- Orengo, C. A. *et al.* CATH—a hierarchic classification of protein domain structures. *Structure* **5**, 1093–1108 (1997).
- Mizuguchi, K., Deane, C. M., Blundell, T. L. & Overington, J. P. HOMSTRAD: a database of protein structure alignments for homologous families. *Protein Sci.* **7**, 2469–2471 (1998).
- Holm, L. & Sander, C. DALI/FSSP classification of three-dimensional protein folds. *Nucleic Acids Res.* **25**, 231–234 (1997).
- Hadley, C. & Jones, D. T. A systematic comparison of protein structure classifications SCOP, CATH and FSSP. *Structure* **7**, 1099–1112 (1995).
- Chothia, C. & Finkelstein, A. V. The classification and origins of protein folding patterns. *Ann. Rev. Biochem.* **59**, 1007–1039 (1990).
- Finkelstein, A. V. & Ptitsyn, O. B. Why do globular proteins fit the limited set of folding patterns? *Prog. Biophys. Mol. Biol.* **50**, 171–190 (1987).
- Cohen, F. E., Sternberg, M. J. E. & Taylor, W. R. Analysis and prediction of protein β -sheet structures by a combinatorial approach. *Nature* **285**, 378–382 (1980).
- Cohen, F. E., Sternberg, M. J. E. & Taylor, W. R. Analysis and prediction of the packing of α -helices against a β -sheet in the tertiary structure of globular proteins. *J. Mol. Biol.* **156**, 821–862 (1982).
- Murzin, A. G., Lesk, A. M. & Chothia, C. Principles determining the structure of β -sheet barrels in proteins: I. A theoretical analysis. *J. Mol. Biol.* **236**, 1396–1381 (1994).
- Taylor, W. R. Defining linear segments in protein structure. *J. Mol. Biol.* **310**, 1135–1150 (2001).
- Taylor, W. R. Searching for the ideal forms of proteins. *Biochem. Soc. Trans.* **28**, 264–269 (2000).
- Murzin, A. G. & Finkelstein, A. V. General architecture of the α -helical globule. *J. Mol. Biol.* **204**, 749–769 (1988).
- Taylor, W. R., Jones, D. T. & Green, N. M. A method for α -helical integral membrane protein fold prediction. *Protein Struct. Funct. Genet.* **18**, 281–294 (1994).
- Taylor, W. R. Protein structure domain identification. *Protein Eng.* **12**, 203–216 (1999).
- Jones, D. T., Taylor, W. R. & Thornton, J. M. A new approach to protein fold recognition. *Nature* **358**, 86–89 (1992).
- Chothia, C. & Murzin, A. G. New folds for all- β proteins. *Structure* **1**, 217–222 (1993).
- Taylor, W. R. Protein structure comparison using bipartite graph matching and its application to protein structure classification. *Mol. Cell. Proteom.* advance online publication 4 March 2002 (DOI 10.1074/mcp.T200001-MCP200).

Supplementary Information accompanies the paper on Nature's website (<http://www.nature.com>).

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to the author (e-mail: wtaylor@nimr.mrc.ac.uk).

correction

Transition of Mount Etna lavas from a mantle-plume to an island-arc magmatic source

Pierre Schiano, Roberto Clocchiatti, Luisa Ottolini & Tiziana Busà

Nature **412**, 900–904 (2001).

We acknowledge that the source of the data for major-element composition of sample 27.2 in the Supplementary Information should have been cited as the work of Kamenetsky and Clocchiatti (ref. 1). We also acknowledge that the discussion of our model for the origin of Mount Etna magmatism would have benefited from referencing the model of multi-stage melting contemporaneous with metasomatic enrichment presented in ref. 1. □

- Kamenetsky, V. & Clocchiatti, R. Primitive magmatism of Mt. Etna: insights from mineralogy and melt inclusions. *Earth Planet. Sci. Lett.* **142**, 553–572 (1996).

The discovery channel

Drug discovery — or at least supplying kit to do it with — is big business.

Multimek 96/384

Beckman Coulter www.beckmancoulter.com

Get ahead with high-density pipetting

Users of Beckman Coulter's Multimek liquid handler can now work in 384-well format with the Multimek 96/384 pipetting head. The device automatically loads 384 tips and offers capacities from 1 to 30 μ l. A disposable tip eliminates liquid carry-over and contamination and includes built-in quality control information. It can be used to automate higher density assay formats and plate replication methods.

High-volume microplates

Porvair Sciences www.porvair-sciences.com

For chemists with compound interests

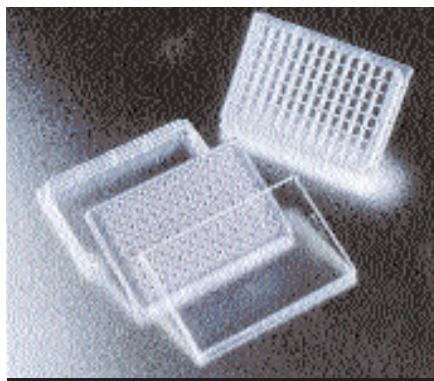
Porvair's new range of high-volume 48- and 96-well microplates with a standard microplate footprint is designed for rapid, large-scale generation of target compounds in combinatorial chemistry. The 96-well plates have a volume of 4 ml, and the 48-well version comes in volumes of 5.0, 7.5 or 10 ml. All have a pierceable seal and are stackable. The plates are also suitable for long-term sample storage.

Falcon HTS 96-multiwell insert

BD Biosciences www.bdbiosciences.com

All of a piece

The BD Falcon HTS 96-multiwell insert system is a cell culture insert assay platform composed of a 1 μ m pore-sized, membrane-based multiwell insert plate, with a media feeder tray and lid. The wells are integrated into a one-piece plate that fits into the feeder tray, suitable for making media changes in one transfer. Individual samples are analysed by transferring the 96-well insert into the BD Falcon 96-square well angled bottom plate. The system can be automated and is suitable for both manual and robotic screening of



Culture clash: the Falcon multiwell insert system.



Mind your fingers: Multimek handling liquid.

compounds in bioavailability, transport, permeability and other cell-based assays.

Biomedical data library

Ardais www.ardais.com

Clinical genomics gets BIGN and better

The Ardais BIGN (biomaterials and information for genomic research) library is a centralized, shared clinical genomics repository for licensed use by qualified biomedical researchers. It contains tissue samples, molecular derivatives and associated clinical information accompanied by databases and bioinformatics tools, and can be used to guide the experimental design of target identification and validation studies. Licensed users gain online access to research-quality tissue samples with in-depth medical, pathology, treatment history and outcomes data, as well as sample formats and molecular derivatives for direct analysis.

BRET² β -Arrestin

Packard BioScience

www.packardbioscience.com

Study the behaviour of orphans

Genes encoding new G protein-coupled receptors (GPCRs) are those for which no natural ligand has been identified — the so-called 'orphan' receptors. The BRET² β -Arrestin Kit monitors GPCR activation and can be used in the identification and characterization of natural and synthetic ligands (agonists and antagonists). This cell-based assay contains the pGFP²- β -arrestin2 fusion construct, which emits a robust BRET² signal when co-expressed with GPCR-Renilla luciferase fusion proteins. Packard BioScience also offers a V₂ vasopressin/ β -arrestin kit for a fully characterized BRET² β -arrestin assay with the V₂ vasopressin receptor.

FLEXstation and HEFP

Molecular Devices www.moleculardevices.com

Three new kits for rapid assaying

FLEXstation calcium and membrane potential assay kits from Molecular Devices are designed for rapid assay development on the FLEXstation development and detection system. Both are already available in a high-throughput format for use on the company's FLIPR system. The HEFP phosphodiesterase assay kit is a high-throughput, non-radioactive assay for screening phosphodiesterases, important drug targets for conditions such as erectile dysfunction and asthma.

MagicMark

Invitrogen

www.invitrogen.com

Weight watcher

MagicMark is a convenient method of estimating protein molecular weight directly on western blots. All eight MagicMark protein bands can be visualized simultaneously with the same antibody conjugate and reagents used to detect a target protein on the blots, using either chemiluminescent or colorimetric methods, thus eliminating the need for extra reagents.

VanKel (VK) 7020

Varian

www.varianinc.com

It dissolves in front of your eyes

This is the first model in Varian's new V-series line of dissolution testers, used to determine the rate at which active ingredients in a dosage form become bioavailable. The VK 7020 features a new shape allowing full visibility and a vessel alignment system. It can be programmed to sample either sequentially or simultaneously, with temperature in each vessel measured and recorded at programmed times. An electric clutch allows the user to start rotating each spindle at a pro-



VanKel 7020: looking for the solution.

grammed time, and a dosage dropper automatically drops tablets into vessels sequentially or simultaneously.

Protein kinases

PanVera

www.panvera.com

Key enzyme in signal transduction

Cyclic AMP-dependent protein kinase (PKA) is a ubiquitous serine/threonine protein kinase that phosphorylates a broad spectrum of substrates. PKA activity regulates cellular processes including cell growth, differentiation, ion-channel conductivity, gene transcription and synaptic neurotransmitter release. The PKA holoenzyme is an inactive tetramer comprising two regulatory and two catalytic subunits. PanVera now markets the cyclic AMP-dependent protein kinase catalytic subunit, of particular interest in drug discovery research. The catalytic subunits are released following the binding of cAMP to the regulatory subunits. The free catalytic subunit is active and does not require cAMP for activity. PanVera's human PKA catalytic subunit has a calculated molecular weight of 43.5 kDa. In the UK and Ireland this kinase is available from Cambridge BioScience.

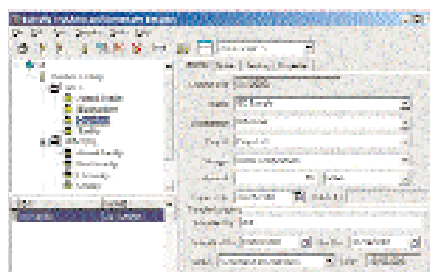
LyoStar II

Kinetics Thermal Systems

www.kineticsgroup.com

Quick-freeze R&D tray

This cascade freeze-drying system provides cooling to -70°C at rates in excess of 2.5°C per minute. The shelf-to-condenser surface ratio ensures ample lyophilization efficiency for even the most aggressive cycles. The software for data monitoring and analysis has been simplified and upgraded. The unit has a shelf temperature stability and accuracy of $\pm 0.5^{\circ}\text{C}$ and features front panel access to the vacuum pump and electrical components, a sealed shelf fluid system to minimize contamination, built-in system diagnostics and caster mounting. A datasheet on the LyoStar II can be downloaded from www.ftsproducts.com/drawing5/



Track star: the Oracle STIS program.

Oracle STIS

ChemSW

www.chemsw.com

Multi-lab sample tracking software

Oracle is ChemSW's new client-server version of the STIS (sample tracking and inventory system) program for monitoring samples and work throughout laboratories, particularly those working in pharmaceuticals and forensics. Users can register samples, create work lists, assign tests, enter test results and check progress and status, with audit trail and reporting. Tests may be assigned to a single sample or to a batch of samples at once. The system includes a suite of functional reports, with customized reports easily created.

Spectrum FT-IR software

PerkinElmer

www.perkinelmer.com

Safe keeping for data

Spectrum Enhanced Security Software that enables 21 CFR Part 11 technical regulatory compliance for the pharmaceutical industry is now available for PerkinElmer Fourier transform-infrared (FT-IR) instruments. The software assists pharmaceutical organizations in meeting the requirements of the Food and Drug Administration's (FDA) 21 CFR Part 11 regulations. The Spectrum Enhanced Security Software meets pharmaceutical companies' need to ensure the integrity of electronic data records that are generated under good laboratory practice (GLP) and good manufacturing practice (GMP) protocols. The system provides tech-

nical compliance in terms of software development and data integrity, with extra security features that help quality assurance and regulatory compliance managers create a secure, reliable environment. Spectrum Enhanced Security Software also provides secure login using a combination of Windows NT and Spectrum features, to ensure that data cannot be generated without permitted access to the system. The software is available for new and existing Spectrum One and Spectrum One NTS FT-IR systems.

Naming process

MedPanel

www.medpanel.com

Name that drug

The new online naming service from MedPanel is designed to smooth the process of naming new drugs and devices. It uses a panel of primary care physicians, working online with the client and a moderator, to generate and then create a shortlist of potential names. The final list is then ranked by each participant in order of preference. MedPanel says that by using this method it has reduced the time required to choose a name for a new drug or device from the usual two or three months to only two or three weeks.

MacVector

Accelrys

www.accelrys.com

Protein analysis for X-ized Mac users

The MacVector program for nucleic acid and protein analysis is now available for use with Apple's OS X operating system. The program is used to edit sequences, analyse PCR primers, search the NCBI's Entrez and BLAST databases for sequences of interest, and perform protein analysis within a single interface. The software includes the Assembly-LIGN module for assembling sequence fragments into a consensus sequence that can be imported into MacVector for analysis.

These notes are compiled in the Nature office from information provided by the manufacturers.

Contacts

Publisher: Fabien Savenay
Editor: Paul Smaglik
Sales Director: Ben Crowe

European Head Office, London

The Macmillan Building
4 Crinan Street
London N1 9XW, UK
Tel +44 (0) 20 7843 4961
Fax +44 (0) 20 7843 4996
e-mail: naturejobs@nature.com

European Manager:

Nevin Bayoumi (4978)

UK/ RoW/ Ireland:

Matt Powell (4953)
Ben Corp (4974)
Andy Douglas (4975)
Laura Pearson (4977)
Netherlands/ Italy/ Iberia:
Evelina Rubio Hakansson (4973)
Scandinavia: Silje Opstrup (4994)

Production Manager:

Billie Franklin
To send materials use London
address above.
Tel +44 (0) 20 7843 4814
Fax +44 (0) 20 7843 4996
e-mail: naturejobs@nature.com

International

Advertising Coordinator:

Hind Berrada (4935)
Naturejobs web development:
Tom Hancock
Naturejobs online production:
Ben Lund

European Satellite Offices

France/ Belgium:

Christine Niox-Chateau
Tel + 33 (0) 1 43 20 16 51
Fax + 33 (0) 1 43 20 51 52
e-mail: c.nioxchateau@nature.com

Germany/ Austria/ Switzerland:

Patrick Phelan/ Kate Turner
Tel + 49 89 54 90 57 11/2
Fax + 49 89 54 90 57 20
e-mails: p.phelan@nature.com
k.turner@nature.com

US Head Office, New York

345 Park Avenue South,
10th Floor, New York, NY 10010-1707
Tel +1 800 989 7718
Fax +1 800 989 7103
e-mail: naturejobs@natureny.com

US Sales Director:

Ben Crowe
US Sales Manager: Peyton Mason

US Advertising Coordinator:

Ashly de Leon

Japan Head Office, Tokyo

MG Ichigaya Building (5F),
19-1 Haraikatanachi,
Shinjuku-ku,
Tokyo 162-0841
Tel +81 3 3267 8751
Fax +81 3 3267 8746
e-mail: k.cowan@naturejpn.com
Japan Manager: Kate Cowan

naturejobs

Reinventing careers

Two centres instrumental in sequencing the human genome are reinventing themselves. In doing so, they are creating new kinds of jobs and, perhaps, different career paths. The Wellcome Trust Sanger Institute (formerly the Sanger Centre) and the Whitehead Institute are vying for talent by structuring advancement and creating scientific work beyond traditional investigator posts. And both offer renewable contracts, rather than exclusively tenure-track positions.

The Sanger is investing £300 million (US\$431 million) to support research in proteomics and functional genomics, while continuing its sequencing work. The Wellcome Trust, which funds the institute in Hinxton, Cambridgeshire, is providing an additional £65 million for information technology infrastructure.

In Cambridge, Massachusetts, the Whitehead Institute is in the early stages of planning a similar switch, perhaps splitting into two units — one that will continue to conduct sequencing and another that will work on post-genomic projects.

The Sanger Institute is competing with industry by offering payment above the Medical Research Council scale. This is similar to the strategy some US foundations have used to attract quality postdocs — and put pressure on the government to make its own stipends more competitive (see *Naturejobs* 3–5; 10 January 2002). As for contracts, Susan Lindquist, director of the Whitehead, believes that more research institutions will move towards using both kinds, giving investigators a choice of flexibility or security.

Finally, both centres anticipate bringing on board PhD biologists to run new facilities, such as microarray centres — like an astronomer running a telescope. Those people will be much more than “high-priced technicians”, Lindquist says. And, if more institutes adopt this model, they will be much in demand.

Paul Smaglik

Naturejobs editor



Contents

REGIONS

Austria's empire-building p4

WWW.NATUREJOBS.COM

Career centre
Information on the
scientific job market

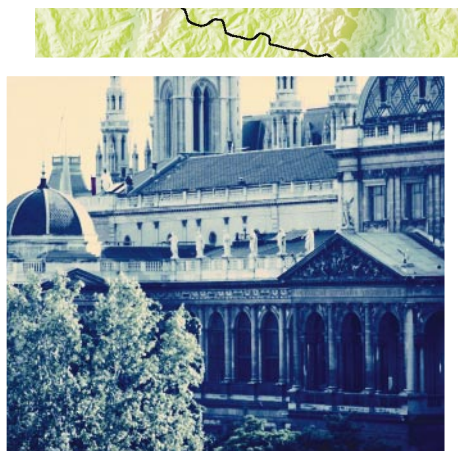
FOCUS

SPOTLIGHT

RECRUITMENT

SCIENTIFIC ANNOUNCEMENTS

SCIENTIFIC EVENTS



Empire building Vienna

After centuries at the hub of central European politics, Vienna is creating a new role at the heart of a fledgling science empire. Three biological institutes are to open there in 2004, creating hundreds of scientific positions. The Institute for Molecular Biotechnology of the Austrian Academy of Sciences (IMBA), the Gregor Mendel Institute for Plant Molecular Biology (GMI), and the Center of Molecular Medicine (CeMM) are also expected to yield pharmaceutical partnerships and biotech spinoffs that will create even more jobs.

The three institutes, all sponsored by the Austrian Academy of Sciences, were conceived when the academy noticed that Austria was “lagging behind” in biology, says Peter Schuster, vice-president of the academy and professor of theoretical chemistry at the University of Vienna. “You have to go beyond the critical mass,” he says, to develop strength in biology.

Reaching that point is neither cheap nor easy, so the academy decided to change the way it creates institutes. Some changes, such as higher salaries and more autonomy for directors and senior scientists, could make recruiting top talent easy. The effects of others, such as complex interactions among institutes, local government and industry, are not so clear.

For example, the IMBA and the GMI will be housed in a building financed by the city of Vienna but connected to the Institute for Molecular Pathology (IMP), which, in turn, is financed by the drug company Boehringer Ingelheim. The IMBA scientists will share equipment, a library, and a cafeteria with those of the IMP. Both the IMBA and the IMP are separate from the university, but scientific discussions are encouraged with the university staff who have lab space in the adjoining Vienna Biocenter.



Josef Penninger says he is excited about building an institute from scratch.

Investment and benefits

For the three biggest drug companies in Vienna, the addition of a new academic research institution to the area means different things. Boehringer Ingelheim will get one of the more obvious benefits — first negotiation on intellectual property generated by the Institute for Molecular Biotechnology. But Boehringer, which is just finishing recruiting chemists, has no plans to increase its Vienna R&D force much beyond the present 175 or so.

Baxter BioScience senses that the new

institutes could help to create a bioscience critical mass. The company is considering building an R&D facility for 300 scientists in Vienna, where it could collaborate with and recruit from the city's university and hospital.

Although Novartis does not plan to increase staff numbers much beyond the 200 currently at its Vienna location, it will benefit from the increased activity. More institutes will mean more symposia, more visiting researchers and a more robust scientific climate.

P.S.

For Josef Penninger, an Austrian researcher recruited from the Amgen Institute in Toronto to head the IMBA, the benefits of this complex arrangement outweigh any uncertainties. “I will be able to pay very good, industrial-type money for really good people,” Penninger says.

Penninger will have enormous freedom to administer the research budget of some 7 million euros (US\$6.1 million). He anticipates that the IMBA



The company does not intend the institute to expand much beyond its current size of 200 scientists.

But the diversity of the neighbouring scientists, rather than their sheer number, is what's important. Many people underestimate the social nature of science, says Nasmyth. He sees the proposed new cafeteria for the university and the three institutes as "the most important thing we will share" — a place where informal meetings may spur interdisciplinary research collaborations.

The presence of Boehringer is another advantage, according to Nasmyth, who believes that Austria could use a more entrepreneurial culture and that academic scientists will benefit from exposure to the business world. "If you're here," he reasons, "you're likely to learn how this whole thing works."

Such relations need not result in a stranglehold on intellectual property. If someone has difficulty in persuading Boehringer that they have a potentially good target or candidate drug, "there's nothing stopping us from going to a third party," Nasmyth says.

That's what happened with Intercell, one of the country's most successful biotech companies. In 1999, former IMP director Max Birnstiel and colleagues from the university of Vienna found a way to use immunology and genetics to develop vaccines. The company, located in an incubator on the Vienna Biocenter's campus, now has three vaccines in clinical trials.

Alexander von Gabain, the company's chief executive, says that Intercell has paved the way for other scientists to secure venture capital funding (once the stock-market picture brightens, that is). He expects that some of the scientists who move into the new facility in 2004 will eventually take that route. ■

Paul Smaglik is editor of *Naturejobs*.

Building on strong foundations: the University of Vienna, far left, and the Institute of Molecular Pathology, left.



Kim Nasmyth believes Vienna will become a major centre for biology research.

R. WALDL

will house about 100 staff — three senior research groups, seven junior groups and technicians — increasing to 150 within a few years. Once they get their general research programmes approved by a scientific advisory board, they will have ample funding and facilities to pursue their specific interests. Unlike university-based colleagues, Penninger says, they will not be burdened by teaching or grant-writing.

Another attraction will be the mouse facility: the IMBA should have room for 10,000–20,000 transgenic rodents. The largest such facility in Austria, this will serve much of central and eastern Europe.

Sharing facilities with the GMI and IMP will be a big plus, says Penninger, because it will create research collaborations. Collaborating with the CeMM could also help translate findings from animals to people.



Dieter Schweizer: combining skills will boost genomics.

Dieter Schweizer, head of cell biology and genetics at the University of Vienna and director of the GMI, agrees. He says that having mouse and plant geneticists working in the same building will be a boon to functional genomics. Access to the signal-transduction and bioinformatics expertise at the IMP will be another plus.

The GMI will be about a third of the size of the IMBA, with three senior groups of 10 researchers each, and four

junior groups of 5–8 each. Although Penninger has not yet named his senior scientists, Schweizer already has three on board. Marjori and Antonius Matzke, who work at an academy institute of molecular biology in Salzburg, will share a post researching epigenetics. Heribert Hirt, currently at the Institute of Microbiology and Genetics at Vienna University, will head a group on plant cell-cycle regulation and signal transduction in the context of host–plant pathogen interactions. Schweizer is still seeking one more senior scientist.

IMP director Kim Nasmyth sees Vienna developing into a strong centre of basic biology research, thanks to the proximity of the three institutes and the university researchers in the Vienna Biocenter. Nasmyth says that the IMP couldn't be the "sole engine" for the area's growth because its own size is limited by Boehringer.

Web links

Vienna Biocenter

♦ www.at.embnet.org/vbc

Institute for Molecular Biotechnology

♦ www.oew.ac.at/english/about/unternehmen/imba.html

Institute of Molecular Pathology

♦ www.imp.univie.ac.at

Center of Molecular Medicine

♦ www.oew.ac.at/english/about/unternehmen/cemm.html

Branching out

Austria is buzzing with activity, and not only in Vienna or in biology.

In Innsbruck, for example, the Austrian Academy of Sciences is considering an institute of quantum optics. The concept — proposed by Anton Zeilinger, an experimental physicist at Vienna University — is timely, because there is already expertise at Innsbruck, and more investment could result

in a world-class centre.

Quantum optics is an area of physics where Austria can compete internationally, as it does not need the costly instruments required by high-energy physics. Zeilinger already gets more postdoc enquiries than he can place.

Other academy plans include a genomics institute in Innsbruck and a mathematics institute in Linz.

P.S.